

DOCUMENT RESUME

ED 292 512

JC 880 147

TITLE College Level Academic Skills Test Review: Report and Recommendations of the Postsecondary Education Planning Commission, 1988. Report 4.

INSTITUTION Florida State Postsecondary Education Planning Commission, Tallahassee.

PUB DATE 25 Feb 88

NOTE 68p.; Prepared in response to Specific Appropriation 537A of the 1987 General Appropriations Act.

PUB TYPE Reports - Research/Technical (143) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC03 Plus Postage.

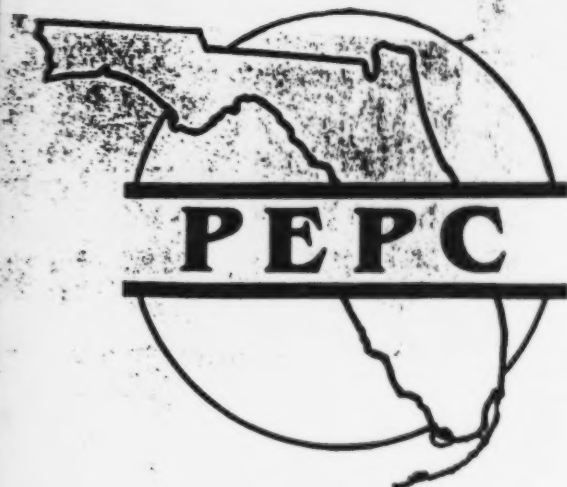
DESCRIPTORS *Achievement Tests; Basic Skills; Community Colleges; Educational Legislation; Higher Education; *Minimum Competency Testing; State Colleges; Statewide Planning; Test Bias; *Test Construction; *Testing Programs; *Test Validity; Two Year Colleges; Writing Evaluation

IDENTIFIERS *College Level Academic Skills Test; *Florida

ABSTRACT

In response to a legislative mandate, this report assesses the validity of the College Level Academic Skills Test (CLAST) and offers recommendations for increasing the passing rate of students. Section I summarizes the Florida State Legislature's interest and intent in establishing CLAST an achievement test--measuring skills in computation, reading, and writing--which must be passed by all students seeking an associate of arts degree at the state's public community colleges and admission into upper division at the state universities. Section II: offers review of the trends and public concerns that contributed to the establishment of the statewide college-level testing program in Florida; highlights testing projects that preceded CLAST; and describes CLAST. Section III presents the results of a study by a consultant team of the psychometric validity of CLAST, revealing that, although most sections of the test were reasonably reliable and valid, improvements in the reliability of the writing test were recommended. Section IV offers recommendations related to the design of CLAST and strategies for improving students' test performance, including: (1) state universities and community colleges should improve linkages with feeder high schools; (2) each institution should track the performance of its students on the college entry-level test to identify at-risk students; and (3) colleges should reinforce essential skills throughout the curriculum rather than in English and mathematics courses alone. CLAST skill statements, codes, and test specifications and the consultants' report are appended. (UCM)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *



ED292512

COLLEGE LEVEL ACADEMIC SKILLS TEST REVIEW

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

W.B. Proctor

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy

Report and Recommendations of the Postsecondary Education Planning Commission

BEST COPY AVAILABLE

1988 - REPORT 4

7C80147

POSTSECONDARY EDUCATION PLANNING COMMISSION

COLLEGE-LEVEL ACADEMIC SKILLS TEST REVIEW

Prepared in Response to
Specific Appropriation 537A
of the
1987 General Appropriations Act

1988 - Report 4

February 25, 1988

TABLE OF CONTENTS

	<u>Page</u>
EXECUTIVE SUMMARY.....	i
I. INTRODUCTION.....	1
II. THE COLLEGE-LEVEL ACADEMIC SKILLS TEST.....	3
Historical Background.....	3
Characteristics of the CLAST.....	7
III. PSYCHOMETRIC VALIDITY OF THE COLLEGE-LEVEL ACADEMIC SKILLS TEST.....	9
IV. ISSUES AND RECOMMENDATIONS.....	13
Design of the CLAST.....	13
Strategies to Improve Performance on the CLAST.....	14
Conclusion.....	18
V. APPENDICES	

EXECUTIVE SUMMARY

The College-Level Academic Skills Test (CLAST) has been required of all students seeking the Associate of Arts degree at the public community colleges and admission into upper division at the state universities since August 1984. In proviso language accompanying Specific Appropriation 537A of the 1987 General Appropriations Act, the Postsecondary Education Planning Commission was directed to:

"conduct a study of the objectives of the College-Level Academic Skills Test. The study shall include, but not be limited to, recommendations on the following: the sufficient number of test items in each form of the test to adequately measure each competency according to acceptable psychometric principles, the difficulty of test items, the adequacy by which test items measure the predetermined competencies and steps and strategies that could be used by postsecondary institutions to increase the passing rate of students. The report shall be submitted on or before March 1, 1988 to the President of the Senate and the Speaker of the House of Representatives."

In September 1987, the Chairman of the Commission directed the Program Committee to produce a report for the full Commission for review and adoption at the February 1988 Commission meeting. Committee meetings were held in October, January, and February, and public testimony was received from various interested parties. The Committee prepared a draft report for public circulation and presentation to the full Commission at its February meeting. The report is to be submitted to the Legislature and the State Board of Education prior to March 1, 1988.

Staff undertook activities in three areas to provide information and analyses to support the Commission in its deliberations.

1. A summary of legislative interest and intent in establishing the CLAST and the Department's implementation of relevant legislative actions was prepared.
2. A team of outside consultants with expertise in test construction in a higher education context was engaged to review the procedures used in the past, to meet with and receive testimony from the offices involved in the construction and monitoring of the test, to review related documents describing the procedures used in developing the test, and to prepare a report assessing the psychometric validity of the test and any modifications deemed necessary.
3. Information was collected from statewide and institutional resources on successful strategies institutions may implement to improve student performance.

In developing background information and analyses for the study, Commission staff met with and received valuable input from the staffs of the State Board of Community Colleges and the Board of Regents, the Office of Student Assessment in the Department of Education, Legislative staff, the Standing

Committee on Student Achievement, and representatives of community colleges, state universities and independent institutions.

The findings and recommendations of the consultant team led by Dr. Ed. Morante, Director of the New Jersey College Outcomes Evaluation Program, are summarized as follows:

1. Statistical analysis of the CLAST indicates that most sections of the test are reasonably reliable and valid. Improvements in the reliability of the writing test are recommended.
2. The psychometrics used in the development and equating of the CLAST are generally appropriate although the test's criteria for development are not always met. In addition, some of the assumptions needed for using the Rasch model of statistics do not appear to have been met.
3. The three verbal sections of the CLAST (Reading, Writing, and Essay) contain items that are at a level of difficulty that should be able to be mastered by all college sophomores regardless of college, degree, or program. Further, the consultants recommended that these verbal sections, with revisions, be considered for use as an entry level basic skills placement test and that a test of higher level skills be developed for college sophomores.
4. The use of an essay component to assess writing as well as the holistic scoring technique are quite appropriate and well implemented. Several improvements in this component are recommended including the combination of the essay with the multiple choice test of writing, and the use of a six-point scale in the holistic scoring process.
5. The Computation section is generally a mathematics test that measures much more than computation (arithmetic). While it contains many items all college sophomores should be able to master, some items appear to be inappropriate for all students, and some items are not mathematics. Revisions are suggested which would strengthen the use of the Computation test to determine proficiency for all sophomores.
6. No evidence was found indicating that the CLAST is biased according to gender or ethnic/racial group.
7. Finally, the impact of the CLAST test has prompted two broad categories of responses. Some educators see current and projected failure rates as argument that the higher standards scheduled for 1989 not be implemented. Others approve of raising standards and accepting the challenge that ensues: improving programs so that more students meet these higher standards. The consultants strongly concur with the latter approach.

Based on the consultants' report and other information obtained throughout the study, the Commission offers the following recommendations concerning the design of the CLAST and strategies designed to improve performance on the test. Several of these recommendations build upon or relate to suggestions contained in the second supplement to the Master Plan, Enhancing the

Participation of Minority and Disadvantaged Students in Postsecondary Education, (March, 1984).

Design of the CLAST

1. The Commission accepts the CLAST consultants' report regarding the psychometric validity and general objectives of the CLAST as well as the absence of any evidence indicating gender or racial/ethnic bias on the test. The Department of Education should thoroughly review all recommendations regarding technical modifications to the CLAST made by the consultants and prepare an assessment of the individual impact of the proposed modifications in terms of student performance on the 1989 standards. This review should be completed no later than August, 1988 to make this information available to the Commissioner's Statewide Panel on CLAST for consideration and action in the Fall of 1988.

Strategies to Improve Performance on the CLAST

2. State universities and community colleges should focus on improved linkages with feeder high schools to assure that potential postsecondary students with special academic needs receive assistance as early as possible.
3. Each institution should track by sub-tests the performance of its students on the college entry-level placement tests to help identify students whose academic preparation may put them at risk of not performing well on the CLAST. In addition, analysis of the performance of students who pass Gordon Rule courses but subsequently fail CLAST is needed to determine if gaps in the curriculum exist.
4. Each institution should develop a strategy to assure that the essential skills defined by the College-Level Academic Skills Program are reinforced throughout the curriculum where appropriate and not solely in the English and mathematics courses.
5. Individual institutions should consider conducting detailed analyses of student performance on the CLAST at the sub-skill level to assist in pinpointing the strengths and weaknesses of the overall lower division curriculum.
6. The State Board of Community Colleges and the Board of Regents should take the lead in determining what individual institutions are doing that result in the more successful performance of their minority populations on the CLAST.
7. In addition to continued basic institutional support, the State should develop guidelines for special program funds to be linked to improved passage rates of CLAST by minority students. The funding should include consideration of the demographic mix

of the respective population served by the institution, the number and percentage of minority and disadvantaged students who attempt and who pass CLAST as well as improvement over time, and other factors such as completion of program requirements and transfer to other postsecondary institutions.

8. Extended periods of time should be made available on CLAST retake administrations for students who indicate a need for more time.
9. Results of the College-Level Academic Skills Test should include analysis of the performance of students with disabilities to determine the effectiveness of existing procedures and the need for any additional accommodations or modifications.
10. Site monitoring by the Department of Education in connection with implementation of the Educational Equity Act should include a review of statewide testing procedures to assure institutional consistency in test administration and the provision of any accommodations for students with disabilities.
11. The commendable efforts to share CLAST materials initiated by the community colleges should be continued and expanded to the state universities and independent institutions willing to assist with the cost of materials duplication and dissemination.
12. Individual institutions should initiate or reinforce existing efforts to inform students of available courses, instructional materials, and other forms of assistance designed to provide the skills measured by the CLAST.

Since its inception the College-Level Academic Skills Test (CLAST) has generated a tremendous amount of interest and effort at both the state and institutional level on the part of students, faculty, administrators, and policy makers to place a renewed emphasis on the importance of general education. While the construction and objectives of the test are basically sound, attention must continue to be focused on the relevance of both the skills tested and the instrumentation of the test itself. The Commission continues to believe that performance will rise to meet expectations and that diminished standards do no service to those who may initially experience difficulty with the test. The information and recommendations contained in this report are intended to strengthen student performance on CLAST as well as the test itself.

I. INTRODUCTION

In proviso language accompanying Specific Appropriation 537A of the 1987 General Appropriations Act, the Postsecondary Education Planning Commission was directed to:

"conduct a study of the objectives of the College-Level Academic Skills Test. The study shall include, but not be limited to, recommendations on the following: the sufficient number of test items in each form of the test to adequately measure each competency according to acceptable psychometric principles, the difficulty of test items, the adequacy by which test items measure the predetermined competencies and steps and strategies that could be used by postsecondary institutions to increase the passing rate of students. The report shall be submitted on or before March 1, 1988 to the President of the Senate and the Speaker of the House of Representatives."

The College-level Academic Skills Test (CLAST) has been required of all students seeking the Associate of Arts degree at the public community colleges and admission into upper division at the state universities since August 1984. In 1983, the Commissioner of Education appointed a statewide panel to establish standards for the test. In its final report, the panel stated that it:

"recognizes that the College-Level Academic Skills Test has been developed as a part of an overall effort in Florida to address the problems of decline in student achievement at all levels of education. The Panel believes that the judicious use of CLAST standards can help reverse the decline in achievement of communication and computation skills and, by doing so, contribute to raising achievement in all other fields in which instruction depends upon communication and/or computation skills. The Panel is convinced that the decline in student achievement is closely associated with the relaxation of graduation requirements in secondary schools as well as in colleges and universities. The consequences of the relaxation of graduation requirements have been further magnified by an attendant relaxation of course grading standards, popularly known as grade inflation."

The passing scores were to be phased in over three periods of time starting in 1984 with scores being raised in 1986 and once again in 1989.

In the four years that the CLAST has been required of college students in Florida a pattern has been established of passing rates being lower when new standards have been applied but then rising over time. The greatest gains have been for minorities even though their scores remain lower than those of whites. Under the 1984 standards, passing rates for whites increased from 92.7 to 93.8 percent, for blacks from 67.3 to 72.5, and for Hispanics from 79.1 to 84.4. Passing scores were raised in 1986. The two administrations that have occurred under the new standards show an increase from 88.9 to 90.7 percent passing for whites, from 51.0 to 58.3 for blacks and a decline for Hispanics from 75.2 to 73.5.

In 1989, the third adjustment of the standards is to occur. At the current levels of student performance, it is estimated that as few as 50 percent of the students will pass the test. However, it is also assumed that reforms in the secondary and college curricula written the past five years will have improved student preparation by that time thus attenuating to some degree the effect of the higher standards. Nevertheless, given the high possible failure rates, a renewed discussion of the purpose and effectiveness of the CLAST has been generated within the postsecondary education community in Florida.

In September 1987, the Chairman of the Commission directed the Program Committee to produce a report for the full Commission for review and adoption at the February 1988 Commission meeting. Committee meetings were held in October, January, and February, and public testimony was received from various interested parties. The Committee has prepared a draft report for public circulation and presentation to the full Commission at its February meeting. The report is to be submitted to the Legislature and the State Board of Education prior to March 1, 1988.

Staff undertook activities in three areas to provide information and analyses to support the Commission in its deliberations.

1. A summary of legislative interest and intent in establishing the CLAST and the Department's implementation of relevant legislative actions was prepared.
2. A team of outside consultants with expertise in test construction in a higher education context was engaged to review the procedures used in the past, to meet with and receive testimony from the offices involved in the construction and monitoring of the test, to review related documents describing the procedures used in developing the test, and to prepare a report assessing the psychometric validity of the test and any modifications deemed necessary.
3. Information was collected from statewide and institutional resources on successful strategies institutions may implement to improve student performance.

In developing background information and analyses for the study, Commission staff met with and received valuable input from the staffs of the State Board of Community Colleges and the Board of Regents, the Office of Student Assessment in the Department of Education, Legislative staff, the Standing Committee on Student Achievement, and representatives of community colleges, state universities and independent institutions.

II. THE COLLEGE-LEVEL ACADEMIC SKILLS TEST

Historical Background

During the early seventies, several trends and public concerns emerged that contributed to the beginning of the statewide college-level testing program in Florida. First, as a reaction to the often flexible and sometimes loosely defined curriculum of the 1960's, a general call for new standards for educational achievement began to emerge. Second, with the explosion of the community college system and the growth of the state university system from three to nine institutions, access to higher education increased rapidly. Third, the concern for standards at the public school level, resulted in an elaborate and comprehensive testing program. And finally, the performance of students on the statewide tests raised concerns about the level of college preparation which resulted in the Legislature requiring students to pass either an ACT or SAT test to be admitted to the State's college of education.

As more students entered higher education with various levels of preparedness, concern over standards began to materialize. In the midst of nationwide debates over the ability of college students to communicate effectively and perform tasks involving basic computation, Florida, in the mid-1970's, began exploring competency-based approaches to education which required less time for completion. Before any consideration of a competency-based program could be made, the specific learning outcomes of a general education would have to be defined and methods would have to be developed for assessing those skills.

The Essential Academic Skills Project

In 1979, the State Legislature took the first formal steps toward the development of the statewide competency-based education and testing program. In that year, as a result of widespread concern that students were graduating from college without attaining at least minimal competency in the communication and computation skills normally associated with a college degree, the Legislature directed the State Board of Education to adopt minimum standards for college-level communication and computation skills. As a response to the legislative directive, the Articulation Coordinating Committee, acting under the authority of the State Board of Education, set out to involve faculty in the process of defining and maintaining academic standards through the establishment of the Essential Academic Skills Project (EASP).

The goals of the EASP were twofold; first, to assure that students entering their freshman year were correctly placed so that they could acquire essential college-level communications and computational skills; and, second, that students entering their junior year had acquired those skills essential to success in college. In order to best accomplish these goals, the EASP was made up of four key component groups all working together. The four groups were (1) the project Executive Committee and the Director, (2) the Statewide Task Force which worked to identify the essential skills, (3) the Institutional Task Forces which assisted in obtaining a consensus model of competencies considered essential to student success in college by reviewing and responding to the lists of skills compiled by the State Task Force, and (4) the Standing Committee on Student Achievement which worked to identify

tests and other means of measuring proficiency in the identified essential skills.

Within the limited time allotted for this first phase of the Essential Academic Skills Project, the Statewide Task Force narrowed the scope of its work to identifying those communications and computational skills which appeared essential for students to possess as they completed their first two years of college and prepared to enter a baccalaureate program. In addition, the Statewide Task Force with the encouragement of the Standing Committee on Student Achievement also included suggestions on methods which could be employed to measure the competencies. The Statewide Task Force identified the first set of communications and computational skills for college students and submitted it for review by the faculties through the Institutional Task Forces in February 1980. The task force coordinators in the community colleges and universities returned over 2,600 surveys to the Statewide Task Force; over one thousand of them contained written comments. A revised list of competencies was once again circulated in survey form in September 1980 after the suggestions, changes and additions from the faculty statewide had been incorporated. By December 1980, a final list of competencies had been completed based on the results of the second institutional survey and the best professional judgement of the State Task Force.

The Statewide Task Force attempted throughout the process of building the consensus model for communication and computation competencies to meet criteria in competence theory and practice specified by a number of related professional organizations including the National Council of Teachers of English College Composition and Communication Conference, the Mathematical Association of America and the National Council of Teachers of Mathematics. As a result, the competencies reflect the following characteristics:

1. The statements are sufficiently specific to indicate (a) what faculties are expected to include in their teaching, (b) what students will be expected to know and be able to do, (c) what skills can be measured to assess the competence--that is, the capability--which is greater than the sum of the individual skills.
2. The statements are logically consistent across the levels of definition.
3. The competencies are generalizable beyond the discipline in which they are learned.
4. The skills and competencies to which they lead are measurable.

As the first phase of the project reached a conclusion, concerns over the quality of teaching and learning in high schools and college continued to fuel the State's efforts to improve educational outcomes. Debates over stricter admission requirements for the State's teacher preparation programs and colleges of education in general, as well as general poor student performance in what many considered basic skill areas brought a renewed call from the Legislature to the State Board of Education for the identification of skills and competencies that students should have to be successful in college. Policymakers were also becoming increasingly interested in a statewide testing program to be used as a requirement for admission to upper division programs.

Legislative concerns over the pejorative effect of grade inflation also generated increasing interest in a testing program as a form of external validation of student academic achievement. In response to these developments, the Essential Academic Skills Project evolved into the College-Level Academic Skills Project in order to expand and further refine the work already underway.

The College-Level Academic Skills Project

The College-Level Academic Skills Project is an organization established in 1981 on a continuing basis by the Commissioner of Education to assist with substantive decisions in the development, implementation and maintenance of the sophomore testing program. The organization is composed of three faculty groups; the Task Force on Communication Skills; the Task Force on Computation Skills; and finally, the Standing Committee on Student Achievement which oversees the testing program and is concerned with the utilization of test results for the improvement of instruction.

The initial task facing the organization was the identification of the skills and competencies required of successful college students. As a result of the previous work under the Essential Academic Skills Project, a consensus had been reached toward a final listing of 117 competencies. After review and approval by the Articulation Coordinating Committee, the list was submitted through the Commissioner of Education, to the State Board of Education for adoption. In September, 1981, the State Board adopted the list of skills as Rule 6A-10.031 of the Florida Administrative Code. Although provision exists for a biennial review of the skills, the list of skills adopted in 1981 remains as the basis for the development of the CLAST. Through the work of the expanded task forces in 1985, minor changes and further refinements were incorporated and ratified by faculty survey across all disciplines, in all participating public and private institutions in the same manner as the original list was created.

In 1981, Legislative concerns over the requirements for admission to colleges of education and more specifically teacher preparation programs and methods for the improvement of education in general, had increased and generated a debate centered around the need for type of sophomore test for admission to upper level study. The Standing Committee on Student Achievement had determined in 1980, after an exhaustive search of commercially available tests, that no existing test could adequately measure student attainment of the complete set of skills the faculty had agreed upon. As a result, the State Board of Education authorized the development of a testing program designed to measure the specific skills delineated by the faculty task forces. In the early stages during field testing and validation, the CLAST could be used for a variety of diagnostic and guidance purposes. Although discussion of the purpose of the statewide testing program was ongoing, no formal decision to implement the test as a requirement for upper division admission had been made at this point.

At the time the State Board adopted the list of college-level communications and computation skills, it further directed the Articulation Coordinating Committee of the DOE (Rule 6A-10.311, Florida Administrative Code) to develop item specifications which would be used in the development of a test to

measure student attainment of the skills. Operating under the supervision of the College-Level Academic Skills Project, the chairpersons of the faculty task forces on mathematics and communications began the process of drafting item specifications and procedures for assessing the skills. After a thorough review process, the respective task forces submitted their item specifications to the Department of Education for the measurement of the computation, reading and writing skills in 1983. Invitations to bid on the development of test items were issued to all of the state community colleges and universities with the stipulation that development teams be composed of faculty involved in the teaching of courses containing the agreed upon skills. The contracts for the initial development of items and all since have provided for multi-institutional teams composed of content and measurement specialists. As test items are developed, they are submitted to the Office of Student Assessment for field testing and analysis before they are actively used as scored items.

The College-Level Academic Skills Test

By October 1982, the Florida Legislature had provided funding and statutory authorization for the initial implementation of the College-Level Testing Program (Section 229.551(3)(j), F.S.). In addition, the 1982 Legislature required that after the Fall term of 1982, all students present CLAST scores for the awarding of an Associate in Arts degree or for admission to upper division status (Chapter 82-180, L.F.). As a result, the CLASP developed specifications for the CLAST involving four subtests. Three of the subtests contain objective items covering reading, computation and writing. The fourth area requires a comprehensive writing sample through a response to one of two essay topics. Until August of 1984, students could satisfy the Legislature's CLAST requirement by simply submitting scores for each of the four subtests. After that date, standards were imposed by the State Board, the attainment of which were required for admission into the upper division and the granting of an Associate in Arts degree.

The setting of standard passing scores was viewed as an important process requiring the same widespread participation as was elicited in the development of the original list of competencies. The Department of Education established a procedure through which performance standards would be proposed by faculty members and other interested individuals who would be impacted by the standards. The procedure involved the identification of standards derived from judgements concerning the level of attainment of the skills expected of all college sophomores, regardless of discipline, in Florida. Panels in all of the state community colleges and universities were then convened for the purpose of suggesting passing scores for each of the four subtests contained in the CLAST. The institutional panels were composed of faculty, administrators, trustees, students and representatives of the public. In much the same manner as was involved in the identification of the initial list of competencies, each institutional panel proposed scores in a process that included taking the CLAST themselves, examining data concerning the performance of students who had taken the test in its earlier administrations and a careful examination of the skills being tested. Institutional panel scores were then forwarded to a statewide panel charged with recommending passing scores on the CLAST to the State Board of Education after evaluating the information available and giving careful attention to the recommendations

of their institutional counterparts. Working from the state level, the panel reached its own decision as to the performance level that should be required of students by the end of their sophomore year. After some consideration, the panel found it necessary to review its position because of the possible impact of the proposed standards upon students entering the program.

After review of the data recorded for the thirty thousand examinees during the first two administrations of the CLAST, the statewide panel, rather than revise its performance expectations, recommended that the imposition of what it considered minimum passing scores be phased in for full implementation in August, 1989. The statewide panel reasoned at the time that such a delay would allow students the full benefit of curricular reforms currently underway in secondary and postsecondary institutions. In order to address the urgent need for some performance standards, the statewide panel recommended to the State Board of Education that interim performance standards be imposed beginning on August 1, 1984 and raised on August 1, 1986 to a higher set of scores that it believed would be attainable by that date. In March 1984, the State Board considered the statewide panel's recommendations and adopted them as Rule 6A-10.312(1) of the Florida Administrative Code. Currently, concern over the possibly high failure rates that would occur under the approaching standards due for implementation in 1989 led to the creation of a statewide review panel by the Commissioner of Education in 1987. This new statewide panel is reviewing the issues related to the implementation of the higher standards for student performance on the CLAST and are expected to make recommendations to the Commissioner of Education and the State Board of Education concerning the standards in the Fall of 1988.

Characteristics of the CLAST

The CLAST is an achievement test that measures skills in computation, reading, and writing. It is administered during the fall, spring, and summer of each year in all community colleges and public universities.

As previously discussed, the design of CLAST was formulated by the College-Level Academic Skills Project (CLASP). CLASP assisted in identifying the skills to be tested and the manner in which they would be tested. It was determined that the computation and reading skills would be assessed using multiple-choice items and that the writing skills would be tested by both multiple-choice items and by requiring the student to compose an essay. All of the writing skills are tested on the essay subtest, but only a portion can be assessed by multiple-choice items.

The passing scores or "standards" for the computation, reading and writing subtests are represented by scale scores. A statistical process is used to convert the number of items correct on each form of the test to a common scale.

A student's essay score is obtained by adding the scores of two raters. Each rater scores an essay on a scale of one to four, so the student's final scores can range from two to eight.

For the first two years of administering CLAST, test results were used only for evaluating curricula and counseling students. As of August 1, 1984,

minimum CLAST scores are required for the award of an associate in arts degree and for admission to upper division in a state university in Florida. However, if an examinee meets the standards on only three of the four subtests, he/she may take up to an additional 36 semester hours in upper division courses before being required to meet the standards on the fourth subtest.

The standards were set at three levels to be implemented in August 1984, August 1986, and August 1989. Hence, the three sets of standards are referred to as the 1984 standards, 1986 standards, and 1989 standards. A student must meet the standards which are in effect at the time the student initially takes CLAST.

Skill Statements

The skills measured by CLAST were identified by the faculties of community colleges and state universities through the CLASP. An initial set was developed in 1981. The skills are to be reviewed by the Articulation Coordinating Committee every two years to ensure that appropriate modifications are made as curricula change. As a result of the skill review process in 1985, a revised set of skills was adopted. The revised set of skills were first tested during the September 1987 administration of CLAST. The equating process (the Rasch model) will compensate for any difference in difficulty of the revised set. (See Appendix A for a list of the skills.)

Item Development Process

Item development for CLAST is conducted by contracts with community colleges and universities. Faculty writing test items follow the guidelines provided in the item specifications, which were also developed with the assistance of faculty. The item specifications describe exactly how each of the items are to be constructed for each skill. The item specifications are available from the CLASP office.

The item-writing contractor is also responsible for conducting an extensive review of each item. This review consists of pilot testing the item on a small group of students and asking students questions about their understanding of each item. With the assistance of a statewide panel, the items are then revised and submitted to the Department.

Each item is then included on the CLAST, although it is not counted as a scored item. This produces statistical data which are used to review each item again. If the item produces acceptable statistics, it can then be included as a scored item on CLAST.

III. PSYCHOMETRIC VALIDITY OF THE COLLEGE-LEVEL ACADEMIC SKILLS TEST

In the summer of 1987, the Commission engaged a team of consultants to undertake an examination of the CLAST to determine if its construction and content were valid according to accepted psychometric principles. A lead consultant, Dr. Edward Morante, was selected who had policy, administrative and technical experience in state-mandated testing at the postsecondary level. Dr. Morante is currently Director of the College Outcomes Evaluation Program for the New Jersey Department of Higher Education, and is past Director of the New Jersey Basic Skills Assessment Program. He holds an Ed.D. in Counseling Psychology from Teachers College, Columbia University. The consultant team also included professors of mathematics and English with experience in statewide testing programs and persons with technical expertise in psychometrics, testing and measurement. The consultant team included representation of minorities and women. (See Appendix B.)

The consultants were requested to "prepare a report on the objectives of the CLAST including but not limited to an analysis of the following: the sufficient number of items in each form of the test to adequately measure each competency according to acceptable psychometric principles, the difficulty of test items, and the adequacy by which test items measure the predetermined competencies."

In preparing this report, the consultant's work and analysis consisted of several key components:

1. Review of background documents and materials related to CLAST. These included:
 - reports on the test administration
 - reports on the test results
 - institutional recommendations and reports
 - reports from various committees and task forces.
2. Interviews with key people involved with CLAST including:
 - current and former Department of Education leaders and test administrators
 - current Commission staff
 - Legislative staff
 - Florida community college and university faculty members and administrators
 - CLAST developers
 - CLAST committee members

A full listing of the individuals that were interviewed and their titles are included in the complete report of the consultants which may be found in Appendix B.

3. Review and analysis of actual CLAST forms and questions from the following test administrations:

September, 1986

March, 1987

June, 1987

September, 1987

4. Observation of an actual CLAST holistic essay reading and table leaders meeting.
5. Review and analysis of extensive item and test results from the above CLAST administrations including:
 - reliability information
 - Rasch item and test data
 - information on bias
 - scatter plots
 - various essay data

The consultant team's report consists of, first, a summary of interviews with selected individuals in Florida involved with the policy formation and planning, implementation, administration and monitoring of the CLAST. Second, the report includes an appraisal of the validity of each of the four subtests - Reading, Writing, Essay and Computation - including a psychometric analysis of the Reading, Writing and Computation subtests and a content analysis of both the verbal and computation sections. The report concludes with a general finding that "the CLAST system produces carefully developed tests with sound psychometric characteristics." More detailed and specific recommendations are offered as a result of the content analysis of the verbal and computation sections.

The findings and recommendations of the consultants are summarized as follows:

1. Statistical analysis of the CLAST indicates that most sections of the text are reasonably reliable and valid. Improvements in the reliability of the writing test are recommended.
2. The psychometrics used in the development and equating of the CLAST are generally appropriate although the test's criteria for development are not always met. In addition, some of the assumptions needed for using the Rasch model of statistics do not appear to have been met.
3. The three verbal sections of the CLAST (Reading, Writing, and Essay) contain items that are at a level of difficulty that should be able to be mastered by all college sophomores regardless of college, degree, or program. Further, it is recommended that these verbal sections, with revisions, be considered for use as an entry level basic skills placement test and that a test of higher level skills be developed for college sophomores.
4. The use of an essay component to assess writing as well as the holistic scoring technique are quite appropriate and well implemented. Several

improvements in this component are recommended including the combination of the essay with the multiple choice test of writing, and the use of a six-point scale in the holistic scoring process.

5. The Computation section is generally a mathematics test that measures much more than computation (arithmetic). While it contains many items all college sophomores should be able to master, some items appear to be inappropriate for all students, and some items are not mathematics. Revisions are suggested which would strengthen the use of the Computation test to determine proficiency for all sophomores.
6. No evidence was found indicating that the CLAST is biased according to gender or ethnic/racial group.
7. Finally, the impact of the CLAST test has prompted two broad categories of responses. Some educators see current and projected failure rates as argument that the higher standards scheduled for 1989 not be implemented. Others approve of raising standards and accepting the challenge that ensues: improving programs so that more students meet these higher standards. The consultants strongly concur with the latter approach.

The consultants' report closed with the following remarks.

The State of Florida has made a significant contribution to assessment in higher education with the implementation of the College-Level Academic Skills Test. Its efforts should be applauded, particularly since the endeavor moves into uncharted areas of performance-based testing on a statewide level.

A part of the on-going concern over the performance standards as established by the state has to do with impact the current and proposed standards may have on particular segments of the higher education population. It is apparent that those Black, Hispanic, Native American, and English as a Second Language students who do take the CLAST consistently do not perform well.

While certain groups do not perform as required, it is crucial that the State remain resolute with the application of the standards. Simultaneously, the State should provide maximum opportunities for all students to perform successfully on the test, and ensure equity in the process. The State should consider a host of academic intervention activities to this end. Such intervention would necessarily include basic skills and developmental education programs, counseling classes and programs, comprehensive learning assistance and tutoring programs, and courses and classes designed to improve basic skills.

Students should have these programs made available at entry into college, before taking college-level courses and the CLAST, and having to experience failure and a disruption of studies.

A well designed and maintained basic skills program would provide a solid academic foundation for all students, but particularly for those who need to overcome histories of educational disadvantage. A program of this

nature would ensure that all students have the opportunity to achieve at minimum standards, but, more importantly, at the standard and level of achievement representative of college level.

As well, the State needs to give serious consideration to the notion of teacher accountability for student success at all levels of the educational system, and move to provide methods for development and implementation of such.

The challenge facing Florida, and other states as well, is in providing both a quality and an equitable system of education. The establishment of standards and a program of assessment is an important first step.

IV. ISSUES AND RECOMMENDATIONS

Design of the CLAST

The consultants' review regarding the psychometric validity of CLAST resulted in a general endorsement of the objectives and construction of the test with specific suggestions for improvement. The consultants further found no evidence of gender or racial/ethnic bias in the test.

The specific recommendations of the consultants are discussed in both the preceding section of this report and in Appendix B. There are several major modifications which require highlighting.

In the computation area the consultant team recommended deletion of the entire Logical Reasoning category from this section of the test. In addition, they suggest deletion of a total of five skills related to geometry and measurement, algebra, arithmetic and statistics. They also suggest revisions to several skill specifications in the areas of arithmetic, algebra and probability. Finally, the addition of skills related to percentage calculation, simultaneous linear equations, and translation of a verbal expression into symbolic form were recommended.

With regard to communication, suggestions included combining the Essay and Writing sections of the test to improve the equation procedure and provide a more accurate picture of students' writing, and increasing the holistic scoring scale to six points to allow for greater differentiation of scores and increase the accuracy of measurement. On the Reading Section, the consultants recommended raising the standard of performance to reflect college-level standards.

The Commission is not in a position to judge in detail the merits of the technical refinements and modifications recommended by the consultants. However, the necessary expertise to make such judgements presently is available through the Office of Student Assessment in the Department of Education and the various panels and advisory groups in place to assist with the design and administration of the CLAST. In assessing the appropriateness of the consultants' recommendations the Department should, to the extent possible, project the individual impact of any proposed modifications in terms of student performance on the 1989 standards. For example, how would deletion of Logical Reasoning or establishment of a six point writing scale impact on student performance based on the 1989 standards? This review should be completed as soon as possible but no later than August 1988 in order to provide this information to the Commissioner's Statewide Panel on CLAST for consideration and action in the Fall of 1988.

Recommendations:

1. The Commission accepts the CLAST consultants' report regarding the psychometric validity and general objectives of the CLAST as well as the absence of any evidence indicating gender or racial/ethnic bias on the test. The Department of Education should thoroughly review all recommendations regarding technical modifications to the CLAST made by the consultants

and prepare an assessment of the individual impact of the proposed modifications in terms of student performance on the 1989 standards. This review should be completed no later than August, 1988 to make this information available to the Commissioner's Statewide Panel on CLAST for consideration and action in the Fall of 1988.

Strategies to Improve Performance on the CLAST

The Commission believes that one of the primary benefits of the College-Level Academic Skills Test is the attention it draws on the part of the collegiate faculty to the communication and computation skills of the lower division students. The College-Level Academic Skills Program and Test were developed by faculty and the continuing improvements the program is designed to encourage are dependent on the involvement of the faculty in diagnosing and remediating the deficiencies identified by the test.

During the course of its review the Commission learned that most, if not all, public and independent colleges and universities in Florida have developed a variety of strategies and materials intended to assist students in achieving the CLAST standards. Rather than describe in detail these individual institutional efforts the Commission makes the following recommendations designed to encourage all higher education institutions to use the CLAST program more extensively and efficiently to achieve higher performance among their students. Several of these recommendations build upon or relate to suggestions contained in the second supplement to the Master Plan, Enhancing the Participation of Minority and Disadvantaged Students in Postsecondary Education, (March 1984).

Better Preparation

Most of the community colleges and state universities in Florida draw the predominant portion of their students from a limited number of high schools. The regional teaching faculties in English and mathematics from both sectors need to meet regularly to discuss the type of skills and levels of preparation needed for college-bound students. The Commission received testimony that even within one service area wide variances in skills levels may exist among high school graduates. Special attention needs to be focused on identifying high school graduates whose academic record indicates they may be at risk of performing poorly on the CLAST. Management information systems technical support at the institutional and systems level may be needed in helping faculty define indicators of risk. The feedback system established pursuant to Section 240.118, F.S., which requires each state university and community college to report to each school district on the performance of its high school graduates, should be useful in this process.

Recommendation:

2. State universities and community colleges should focus on improved linkages with feeder high schools to assure that potential postsecondary students with special academic needs receive assistance as early as possible.

Placement tests administered at entry-level can also assist in identifying students whose academic preparation may put them at risk of not performing well on the CLAST. This is of particular importance for students whose high school academic record may not be as useful because they have been out of high school for a number of years or are from out-of-state. Of particular concern are the students who marginally meet the college entry-level standards especially given the testimony by some faculty that the standards may be too low.

In addition, each college and university needs to analyze the sub-population of those who fail the CLAST to determine if it can identify earlier those at risk. Indicators of risk specific to each institution need to be developed, collected and analyzed. Faculty members who teach students in the required Gordon Rule courses need to scrutinize the performance of these students in the classroom and their subsequent performance on the CLAST to determine if there are any gaps in the curriculum which resulted in the anomaly of students passing the Gordon Rule courses but failing the CLAST.

Recommendation:

3. Each institution should track by sub-tests the performance of its students on the college entry-level placement tests to help identify students whose academic preparation may put them at risk of not performing well on the CLAST. In addition, analysis of the performance of students who pass Gordon Rule courses but subsequently fail CLAST is needed to determine if gaps in the curriculum exist.

Curricular Improvement

Many institutions have adopted a strategy of assuring that the essential skills defined by the College-Level Academic Skills Program are supported throughout the curriculum and not solely in the English and mathematics courses. For example, greater use of essay examinations rather than multiple choice or short answer tests provide students with greater opportunity to practice writing skills. The faculty assessment of such examinations should insist on competent levels of written English. Such an approach not only offers more practice for the student but reinforces the local applicability of these skills in the student's academic pursuits.

Recommendation:

4. Each institution should develop a strategy to assure that the essential skills defined by the College-Level Academic Skills Program are reinforced throughout the curriculum where appropriate and not solely in the English and mathematics courses.

Detailed analysis of student performance on the CLAST at the sub-skill level for each institution would assist in pinpointing the strengths and weaknesses of the overall lower division curriculum. The success of such an effort would depend upon the active involvement of the institutional faculty. St. Petersburg Junior College has developed a model in this area which may be

copied or modified by other institutions. The St. Petersburg model's strength lies in its detailed analysis of student performance at the sub-skill level and the relating of the sub-skills to the curriculum. The information on student performance is relayed back to the instructors in a form that identifies areas of low achievement to make necessary adjustments in their instruction. Each student's performance is analyzed by skill and skill area by whether a student passes or fails that subtest.

Recommendation:

5. Individual institutions should consider conducting detailed analyses of student performance on the CLAST at the sub-skill level to assist in pinpointing the strengths and weaknesses of the overall lower division curriculum.

Special Populations

In 1986 the Commission completed an analysis of community college's performance in preparing minority students for the CLAST. The report provides detailed data on performance of minorities by institution over multiple test administrations. Its conclusion suggests that some institutions have been consistently more successful than others in improving minority students' preparation and performance on the CLAST. The State could provide funds to help support institutions that demonstrate a commitment to improving passage rates of CLAST by minority students. The funds should include consideration of the demographic mix of the respective population served by the institution, should address the participation and success rates of minorities on CLAST, and should include a responsibility for dissemination of successful strategies to other institutions.

Recommendation:

6. The State Board of Community Colleges and the Board of Regents should take the lead in determining what individual institutions are doing that result in the more successful performance of their minority populations on the CLAST.
7. In addition to continued basic institutional support, the State should develop guidelines for special program funds to be linked to improved passage rates of CLAST by minority students. The funding should include consideration of the demographic mix of the respective population served by the institution, the number and percentage of minority and disadvantaged students who attempt and who pass CLAST as well as improvement over time, and other factors such as completion of program requirements and transfer to other postsecondary institutions.

Another issue related to successful performance rates on the CLAST is the special concern of students for whom English is not their native or first language as well as some students with disabilities. Studies by the Department of Education and some institutions indicate that the amount of time available for taking the communication subtests may not be adequate for a small percentage of students. The CLAST subtests are not designed to be

completed in a restricted period of time. However, pragmatic considerations such as test security, student fatigue, availability of proctoring and space, require that reasonable limits be placed on the amount of time available for the subtests. Since only small numbers of students are present at retakes of the examination, it would be feasible to allow more extended periods of time at these administrations. The few students who need more time would thus not be penalized by time considerations. There is no penalty for retakes and a student may attempt the test as often as he or she wishes.

For learning disabled and other students with disabilities, testing at the postsecondary level represents not only a means of diagnosis but also a potential challenge to be met upon entrance and as a condition for exit from our public colleges and universities. Although the College-Level Academic Skills Test includes provisions for accommodations for students with disabilities, additional quantitative information is needed to determine the effectiveness of existing procedures and the need for modifications. At present, the results of statewide assessment testing in the public schools are available by disability. A similar breakout is not yet available at the postsecondary level for CLAST.

In 1984, the Florida Legislature enacted the "Florida Educational Equity Act" which requires public educational institutions to develop and implement strategies to increase participation of students of a particular race, national origin, sex, handicap or marital status which have been traditionally underrepresented. The Department of Education's Office of Equal Educational Opportunity is charged with assisting and periodically reviewing educational agencies to assure compliance with the law. This review process represents an opportunity for assuring that statewide testing procedures are being administered consistently for students with disabilities.

Recommendations:

8. Extended periods of time should be made available on CLAST retake administrations for students who indicate a need for more time.
9. Results of the College-Level Academic Skills Test should include analysis of the performance of students with disabilities to determine the effectiveness of existing procedures and the need for any additional accommodations or modifications.
10. Site monitoring by the Department of Education in connection with implementation of the Educational Equity Act should include a review of statewide testing procedures to assure institutional consistency in test administration and the provision of any accommodations for students with disabilities.

Information Sharing Among Institutions

In 1983 the College-Level Academic Skills Program Cooperative was formed through the initiative of the Presidents' Council and the Council on Instructional Affairs to encourage the development and sharing of educational

materials related to CLAST among the 28 community colleges. The Cooperative developed an inventory notebook of instructional materials - testing, audio-visual, computerized handbooks and textbooks - used by the colleges for teaching the skills related to the CLAST. Plans to update this inventory are currently underway. State universities and independent institutions willing to contribute to the cost of this dissemination should be invited to participate.

Recommendation:

11. The commendable efforts to share CLAST materials initiated by the community colleges should be continued and expanded to the state universities and independent institutions willing to assist with the cost of materials duplication and dissemination.
12. Individual institutions should initiate or reinforce existing efforts to inform students of available courses, instructional materials, and other forms of assistance designed to provide the skills measured by the CLAST.

Conclusion

Since its inception the College-Level Academic Skills Test (CLAST) has generated a tremendous amount of interest and effort at both the state and institutional level on the part of students, faculty, administrators, and policy makers to place a renewed emphasis on the importance of general education. While the construction and objectives of the test are basically sound, attention must continue to be focused on the relevance of both the skills tested and the instrumentation of the test itself. The Commission continues to believe that performance will rise to meet expectations and that diminished standards do no service to those who may initially experience difficulty with the test. The information and recommendations contained in this report are intended to strengthen student performance on CLAST as well as the test itself.

APPENDIX A

**CLAST Skill Statements and Codes
and Test Specifications**

**College-Level Academic Skills Test
Skill Statements and Codes and Test Specifications**

**Office of Student Assessment
Florida Department of Education, 1987**

The revised set of skills will be first tested during the September 1987 administration of CLAST.

COMPUTATION SKILLS

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		Arithmetic Skills:
1A1	-	Adds, subtracts, multiplies, and divides positive rational numbers
-	1A1	Adds, subtracts, multiplies, and divides rational numbers
1A2	-	Adds, subtracts, multiplies, and divides positive rational numbers in decimal form
-	1A2	Adds, subtracts, multiplies, and divides rational numbers in decimal form
-	1A3	Calculates percent increase and percent decrease
2A1	2A1	Recognizes the meaning of exponents
2A2	2A2	Recognizes the role of the base number in determining place value in the base-ten numeration system and in systems that are patterned after it
2A3	2A3	Identifies equivalent forms of positive rational numbers involving decimals, percents, and fractions
2A4	2A4	Determines the order-relation between magnitudes
-	2A5	Identifies a reasonable estimate of a sum, average, or product of numbers
3A1	3A1	Infers relations between numbers in general by examining particular number pairs
3A2	3A2	Selects applicable properties for performing arithmetic calculations
4A1	-	Solves real-world problems which do not require the use of variables
-	4A1	Solves real-world problems which do not require the use of variables and which do not involve percent
-	4A2	Solves real-world problems which do not require the use of variables and which do require the use of percent
4A2	4A3	Solves problems that involve the structure and logic of arithmetic
		Algebra Skills:
1C1	1C1	Adds, subtracts, multiplies, and divides real numbers
1C2	1C2	Applies the order-of-operations agreement to computations involving numbers and variables
1C3	1C3	Uses scientific notation in calculations involving very large or very small measurements
1C4	1C4	Solves linear equations and inequalities
1C5	1C5	Uses given formulas to compute results when geometric measurements are not involved
-	1C6	Finds particular values of a function
-	1C7	Factors a quadratic expression
-	1C8	Finds the roots of a quadratic equation
2C1	2C1	Recognizes and uses properties of operations

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		Algebra Skills Continued
2C2	2C2	Determines whether a particular number is among the solutions of a given equation or inequality
2C3	2C3	Recognizes statements and conditions of proportionality and variation
2C4	2C4	Recognizes regions of the coordinate plane which correspond to specific conditions
3C1	3C1	Infers relations among variables
3C2	3C2	Selects applicable properties for solving equations and inequalities
4C1	4C1	Solves real-world problems inviting the use of variables, aside from commonly used geometric formulas
4C2	4C2	Solves problems that involve the structure and logic of algebra
		Geometry and Measurement Skills:
1B1	1B1	Rounds measurements to the nearest given unit of the measuring device
1B2	1B2	Calculates distances, areas, and volumes
2B1	-	Recognizes horizontal, vertical, parallel, perpendicular, and intersecting lines
2B2	2B1	Identifies relationships between angle measures
2B3	2B2	Classifies simple plane figures by recognizing their properties
2B4	2B3	Recognizes similar triangles and their properties
2B5	2B4	Identifies appropriate types of measurement of geometric objects
3B1	3B1	Infers formulas for measuring geometric figures
3B2	3B2	Selects applicable formulas for computing measures of geometric figures
4B1	4B1	Solves real-world problems involving perimeters, areas, and volumes of geometric figures
4B2	4B2	Solves real-world problems involving the Pythagorean property
		Logical Reasoning and Computer Technology Skills:
1E1	1E1	Deduces facts of set-inclusion from a diagram
2E1	2E1	Identifies simple and compound statements and their negations
2E2	2E2	Determines equivalence or nonequivalence of statements
2E3	2E3	Draws logical conclusions from data
2E4	2E4	Recognizes that an argument may not be valid even though its conclusion is true
2E5	-	Distinguishes fallacious arguments from nonfallacious ones
2E6(NT)	-	Recognizes proof by contradiction
3E1	3E1	Infers valid reasoning patterns and expresses them with variables
3E2	3E2	Selects applicable rules for transforming statements without affecting their meaning
4E1	4E1	Draws logical conclusions when facts warrant them
2F1	-	Identifies characteristics of tasks which computers perform well
2F2	-	Identifies the human functions necessary to utilize computers
2F3	-	Identifies possible abuses of computer use

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		<u>Statistics Skills, Including Probability:</u>
101	101	Identifies information contained in bar, line and circle graphs
102	102	Determines the mean, median, and mode of a set of numbers
103	-	Selects the sample space associated with an experiment
-	103	Counts subsets of a given set
201	-	Recognizes the normal curve and its properties
-	201	Recognizes properties and inter-relationships among the mean, median, and mode in a variety of distributions
202	-	Recognizes samples that are representative of a given population
-	202	Chooses the most appropriate procedures for selecting an unbiased sample from a target population
203	203	Identifies the probability of a specific outcome in an experiment
301	301	Infers relations and makes accurate predictions from studying particular cases
401	401	Solves real-world problems involving the normal curve
402	402	Solves real-world problems involving probabilities

READING SKILLS

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		<u>Literal Comprehension Skills:</u>
A1	A1	Recognizes main ideas
A2	A2	Identifies supporting details
A3	A3	Determines meanings of words on the basis of context
		<u>Critical Comprehension Skills:</u>
B1	B1	Recognizes the author's purpose
-	B2	Identifies author's overall organizational pattern
B2	B3	Distinguishes between statement of fact and statement of opinion
B3	B4	Detects bias
B4	B5	Recognizes author's tone
B5	B6	Recognizes explicit and implicit relationships within sentences
B6	B7	Recognizes explicit and implicit relationships between sentences
B7	B8	Recognizes valid arguments
B8	B9	Draws logical inferences and conclusions

WRITING SKILLS

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		<u>Word Choice Skills:</u>
B1a	B1a	Uses words which convey the denotative and connotative meanings required by context
B1b	-	Avoids slang, jargon, cliches, and pretentious expressions
B1c	B1c	Avoids wordiness

<u>Initial</u>	<u>Revised</u>	<u>Skills by Broad-Skill Category</u>
		<u>Sentence Structure Skills:</u>
B2a	B2a	Places modifiers correctly
B2b	B2b	Coordinates and subordinates sentence elements according to their relative importance
B2c	B2c	Uses parallel expressions for parallel ideas
B2d	B2d	Avoids fragments, comma splices, and fused sentences
B3c	-	Avoids awkward constructions
		<u>Grammar, Spelling, Capitalization, and Punctuation Skills:</u>
B4a	B4a	Uses standard verb forms
B4b	B4b	Maintains agreement between subject and verb, and between pronoun and antecedent
B4c	B4c	Uses proper case forms
-	B4e	Uses adjectives and adverbs correctly
B5	B5	Uses standard practices for spelling, punctuation, and capitalization

ESSAY SKILLS

The skills measured on the multiple-choice writing subtest are also tested on the essay subtest.

<u>Initial</u>	<u>Revised</u>	<u>Skills</u>
A1	-	Selects subject which lends itself to expository writing
-	A1	Selects subject which lends itself to development
A2	-	Determines the purpose for writing
-	A2	Determines the purpose and the audience for writing
A3	A3	Limits a subject to a topic which can be developed within the requirements of time, purpose, and audience
A4	-	Formulates a thesis statement which reflects the purpose
-	A4	Formulates a thesis or main idea statement which reflects the purpose and the focus
A5a	A5a	Provides adequate support which reflects the ability to distinguish between generalized and concrete evidence
A5b	-	Arranges the main ideas and supporting details in an organizational pattern appropriate to the expository purpose
-	A5b	Arranges the ideas and supporting details in an organizational pattern appropriate to the purpose and the focus
A5c	-	Writes unified prose in which all supporting material is relevant to the thesis statement
-	A5c	Writes unified prose in which all supporting material is relevant to the thesis or main idea statement
A5d	A5d	Writes coherent prose, providing effective transitional devices which clearly reflect the organizational pattern and the relationships of the parts
-	B1b	Avoids inappropriate use of slang, jargon, cliches, and pretentious expressions
B3a	B3a	Uses a variety of sentence patterns
B3b	B3b	Avoids unnecessary use of passive construction
B4d	B4d	Maintains a consistent point of view
B6	B6	Revises, edits, and proofreads to assure clarity, consistency, and conformity

Test Specifications

The test specifications define the number of scored items and the number of items embedded for developmental purposes. They also show how many CLAST skills are in each broad-skill category and the number of items that will appear on CLAST in each broad-skill category. (The number of items in each broad-skill may vary slightly given the constraints of the items available.)

The test specifications for 1986-87 and 1987-88 follow. The test specifications for 1986-87 are based on the initial set of skills and the specifications for 1987-88 are based on the revised set of skills.

CLAST TEST SPECIFICATIONS October 1986, March 1987, and June 1987

SUBTEST	BROAD-SKILL AREA	SKILLS	NUMBER OF ITEMS		
			Scored	Develop.	Total
Reading	Literal Comprehension	03	10		
	Critical Comprehension	08	26		
	Total	11	36	8	44
Writing	Word Choice	03	08		
	Sentence Structure	05	13		
	Grammar, Spelling, Capitalization & Punct.	04	12		
	Total	12	33	5	40
Computation	Arithmetic	10	10		
	Algebra	13	13		
	Geometry & Measurement	11	11		
	Logical Reasoning	10	09		
	Computer Technology	03	03		
	Statistics, Including Probability	09	09		
	Total	56	56*	9	65
Essay	Two topics will be given. The examinee will choose one on which to write the essay.				

*The computation subtest will include one item per skill for a total of 55 items. One additional item will be included in one of the skill areas as needed to complete the statistical requirements of the test.

CLAST TEST SPECIFICATIONS October 1987, March 1988, and June 1988

SUBTEST	BROAD-SKILL AREA	SKILLS	NUMBER OF ITEMS		
			Scored	Develop.	Total
Reading	Literal Comprehension	03	09		
	Critical Comprehension	09	27		
	Total	12	36	8	44
Writing	Word Choice	02	06		
	Sentence Structure	04	13		
	Grammar, Spelling, Capitalization & Punct.	03	16		
	Total	11	35	5	40
Computation	Arithmetic	13	13		
	Algebra	16	16		
	Geometry & Measurement	10	07		
	Logical Reasoning	08	07		
	Statistics, Including Probability	09	07		
	Total	56	50	9	59
Essay	Two topics will be given. The examinee will choose one on which to write the essay.				

APPENDIX B

Consultants' Report

An Analysis of the
College Level Academic Skills Test
for the Florida Postsecondary
Education Planning Commission
January, 1988

Edward A. Morante, Ed. D.
David Osberg, Ph. D.
Nina Jemmott, M. A.
Charles Pine, Ph. D.
Roman Voronka, Ph. D.

EXECUTIVE SUMMARY

Background

The College Level Academic Skills Test (CLAST) is one of several initiatives in Florida created by the legislature and implemented by the Department of Education to raise educational standards in the state. The CLAST is a test of reading, writing, and mathematics that must be passed in order for students in the state's colleges and universities to receive an associate degree or continue to the upper division of a university.

The CLAST has four components: Reading, Writing, Essay, and Computation. The test, including its specifications, items, procedures, and standards, has been collaboratively designed and implemented by Florida faculty and administrators, testing experts in the Department of Education, and various Florida postsecondary test contractors coordinated by the University of Florida, Gainesville.

The test was first administered in 1982 on a trial basis with full implementation of the "gateway provision" in 1984. Passing standards were raised in 1986 and are scheduled to be raised again in 1989. A statewide panel appointed by the Commissioner of Education is currently examining these standards. In addition, the 1987 legislature directed the Postsecondary Education Planning Commission to conduct a technical analysis of the CLAST.

Purpose and Procedure

The purpose of this report is to examine the "psychometric validity" of the CLAST. This refers both to the statistics used in the construction and verification of the test as well as its content validity.

In carrying out these evaluations, we

- made several visits to Florida;
- interviewed a variety of individuals involved in the development, scoring, and use of the CLAST;
- reviewed numerous background reports and booklets;
- reviewed and analyzed all of the specifications and items on the four most recent forms of the CLAST;
- reviewed and analyzed a wide variety of computer printouts containing statistics and related data on the CLAST.

Conclusions and Recommendations

1. Statistical analysis of the CLAST indicates that most sections of the test are reasonably reliable and valid. We recommend that the reliability of the writing test be improved.

2. The psychometrics used in the development and equating of the CLAST are generally appropriate although the test's criteria for development are not always met. In addition, some of the assumptions needed for using the Rasch model of statistics do not appear to have been met.
3. The three verbal sections of the CLAST (Reading, Writing, and Essay) contain items that are at a level of difficulty that should be able to be mastered by all college sophomores regardless of college, degree, or program.

Further, we recommend that these verbal sections, with revisions, be considered for use as an entry level basic skills placement test and that a test of higher level skills be developed for college sophomores.

4. The use of an essay component to assess writing as well as the holistic scoring technique are quite appropriate and well implemented. We recommend several improvements including the combination of the essay with the multiple choice test of writing and the use of a six-point scale in the holistic scoring.
5. The Computation section is generally a mathematics test that measures much more than computation (arithmetic). While it contains many items all college sophomores should be able to master, some items appear to be inappropriate for all students, and some items are not mathematics. We believe that, with revisions, the Computation test could continue to be used to determine proficiency for all sophomores.
6. No evidence was found indicating that the CLAST is biased according to gender or ethnic/racial group.
7. The impact of the CLAST test has prompted two categories of response. Some educators see current and projected failure rates as argument that the higher standards scheduled for 1989 not be implemented. Others approve of raising standards and accept the challenge that ensues: improving programs so that more students meet these higher standards. We strongly concur with the latter approach.

INTRODUCTION

The State of Florida has, for more than ten years, been actively involved in raising the standards of its public educational system. This effort has focused on both elementary and secondary education as well as post-secondary education. It has included legislation and policies that have called for increased requirements, provided additional funding, and initiated testing of students at various levels.

Among the state's public colleges and universities, students are tested at entry, placed in "college prep" courses (remediation) if deemed appropriate, required to do much writing, and to complete a specified number of courses in English and mathematics (the "Gordon rule"). Students are also required to successfully pass a state-administered standardized examination, the College Level Academic Skills Test (CLAST).

The CLAST is designed to measure college-level communication and computation skills of college students at the end of their sophomore year. It is an achievement test battery consisting of separate tests in four areas: reading, writing, computation, and essay. Raw scores on these tests are converted to scaled scores for the purpose of providing a stable assessment of proficiency across test forms in each of the subject areas. These scaled scores are used to determine whether individual students demonstrate proficiency in basic academic skills sufficient to be permitted to enter the upper division of a university or to receive an Associate in Arts degree. Standards for this level of sufficiency have been established by the State Board of Education on recommendation of an expert panel. Increasing levels of proficiency standards have been established. These levels of proficiency, or cutoff scores, are referred to as the 1984, 1986, and 1989 standards.

The Florida legislature has asked the Postsecondary Education Planning Commission (PEPC) to prepare a report on the CLAST, reviewing the test for psychometric validity. The Commission, in turn, has contracted us as consultants to "prepare a report on the objectives of the CLAST including but not limited to an analysis of the following: the sufficient number of items in each form of the test to adequately measure each competency according to acceptable psychometric principles, the difficulty of test items, and the adequacy by which test items measure the predetermined competencies."

This report addresses the request of the Florida Postsecondary Education Planning Commission.

Procedure

In preparing this report, the work and analysis consisted of several key components:

1. Review of background documents and materials related to CLAST; these included:

- reports on the test administration
- reports on the test results
- institutional recommendations and reports
- reports from various committees and task forces

2. Interviews with key people involved with CLAST including:

- former Department of Education leaders
- current PEPC staff
- legislative analysts
- Florida faculty members and institutional administrators
- Department of Education test experts
- CLAST developers
- CLAST committee members

A full listing of the individuals that were interviewed and their titles are included in Appendix A.

3. Review and analysis of actual CLAST forms and questions from the following test administrations:

September, 1986
 March, 1987
 June, 1987
 September, 1987

4. Observation of an actual CLAST holistic essay reading and table leaders meeting.

5. Review and analysis of extensive computer printouts of item and test results from the 9/86, 3/87, and 6/87 CLAST administrations including:

- reliability information
- Rasch item and test data
- information on bias
- scatter plots
- various essay data

The following report begins with a summary of interviews and discussions held with a variety of key Florida people involved with the CLAST. It then provides a psychometric analysis of the data followed by content validity analyses of the CLAST. The report concludes with observations and recommendations concerning impact.

SUMMARY OF INTERVIEWS

The persons interviewed were given the understanding of the purpose of the interviews and discussions. They were told that their responses and opinions would not be person-specific, but rather would be represented in a summary account of all interviews. Those interviewed were receptive to the process and expressed an open willingness to discuss the CLAST and related issues.

The following represents a synthesis of statements and discussion:

- Those interviewed were clear about the catalyst for the creation of the CLAST. Apparently, concern arose from several sources in the 1970's about the quality of students entering and exiting from Florida's colleges and universities. This concern was part of a variety of similar questions related to K-12 education, vocational education, and articulation between community and junior colleges and four year institutions. As a response to the quality of student learning in public postsecondary institutions, the legislature acted to create an assessment program. The assessment program was to determine a) what the proficiencies of college students were, and b) a method of maintaining or raising standards of proficiencies.
- Those interviewed were clear and satisfied with the idea that, in setting up the program, much care was taken to involve a wide cross-section of Florida Higher Education faculty and administration. It was the faculty who made many of the decisions regarding implementation of what has become known as the College Level Academic Skills Program (CLASP) and the College Level Academic Skills Test (CLAST).
- Although it was not clear to the majority of those interviewed of the details of the actual contracting and constructing of the test, it was clear that it was the decision of the statewide faculty committees to focus on what students should know at the end of their sophomore year, within the guidelines of the legislation, in the areas of reading, writing, and computation. It was also generally known that the faculty committees would have to create their own test as no instrument available at that time was deemed appropriate.
- Those interviewed expressed commitment and pride in the involvement of faculty in determining objectives and standards, as opposed to legislative mandate in determining same. Several methods were used to determine the standards, including surveys completed by college instructors, and ongoing faculty task forces.
- Opinions differed among those interviewed on what contributed more to the standard agreed upon: whether it was faculty and college expectations of what college sophomores should know, or the actual impact of the standards on the sophomore population. This latter consideration is based on the "impact data," which shows failure and passing rates of specific subgroups tested, and the impact those rates have and might have on the student population.

- All those interviewed felt that educational reform was past due in Florida and that the CLAST had helped this reform movement to varying degrees. Only one person felt that the colleges and universities themselves should be expected to bring about improvement and reform without state mandate. All others felt that the state's involvement has served as the catalyst for improvement in the areas of: coalesced thinking in terms of standards; increased communications across the faculties and institutions; changed curricula; improved student achievement; faculty and institutional accountability.
- Many of those interviewed expressed serious concern about individual institutional preoccupation with test scores and results. Apparently, in order to raise student scores, many institutions have created "crash courses" and teach ways to pass the test. Other institutions "screen" students, selecting only those students likely to pass to take the test. These responses are seen as attempts by the institutions to avoid embarrassment of high failure rates.
- There was no consensus on the change in state policy to allow freshmen to take the test; some of those interviewed were optimistic; others not. Some people interviewed were unaware of changes, real or contemplated.
- Responses on the appropriateness of the CLAST standards ranged from reasonable to too low on the verbal sections and from reasonable to too high on the math section. Virtually all of those interviewed felt that the standards should be raised to the 1989 level. One person thought that the legislative intent of the program had been undermined since some of the standards were too low. At least one person thought that the standards measured would be better served at the high school or even elementary school level. One person felt strongly that the low level of skills measured was one of the state's "well kept secrets," and expressed great dissatisfaction with the low standards. There was less agreement on when or how the standards should be raised.
- There was general concern about the impact of the CLAST standards and what raising these standards would lead to. Many contrasted the level of some of the skills measured with the relatively poor performance of many students, especially minority students. There was also some concern about the lower pass rate for CLAST than has been found for college courses.
- All those interviewed expressed varying degrees of concern about the performance of blacks, Hispanics, and non-native Americans. Contributing to the low performance of blacks was inadequate pre-college preparation and the continuing effects of segregation; of Hispanics, non-native Americans, and others was second language and a history of educational disadvantage. There was disagreement over whether the solution was one to be found at the level of higher education or elementary and secondary education.
- Many of those interviewed expressed concern about the overall standards of a college education in the state. Not a few suggested that funding

based on Full-Time-Enrollment had impact not only on standards, but also on college curriculum, faculty, expectations of a college student, accountability, and performance on CLAST.

- Most of those interviewed were concerned with some specific aspect of the CLAST. The concerns ranged from the selection and difficulty of items on the Computation section of the test: whether non-math related majors would ever need some of the skills tested; to the selection and facility of items on the Reading and Writing sections of the test; to improving the essay exam by moving to a six point scale, combining it with the writing test, and dropping the refereeing of many papers.
- A few of those interviewed expressed some concern that the CLAST, as a program developed by faculty, with skills determined by faculty, will have a reverse effect on the issue of faculty accountability, and will lead to faculty placing the responsibility to achieve on the students alone, and not on both.
- All those interviewed felt that, in spite of the low pass rates for minorities, and in spite of disadvantages and difficulties in implementation, the CLAST and CLASP have had and continue to have serious and positive implications for secondary schools and higher education in Florida, and should be continued as a program and a policy. They also felt that the test and the program should be changed and improved, as appropriate and necessary.

PSYCHOMETRIC ANALYSIS

Overview

The purpose of this review was to assess the psychometric characteristics of three of the four CLAST tests: Reading, Writing, and Computation. The primary focus of the review was the three forms administered in September of 1986 and March and June of 1987. However, available data from previous years were also reviewed in order to provide a more comprehensive background for reporting related to measurement-related aspects of the tests. All data reviewed were supplied by officials of the College Level Academic Skills Project (CLASP).

Validity

A test is said to have validity to the degree to which it adequately measures the specific skills or traits that it was designed to measure. One must be confident of a test's validity if one is to have confidence on decisions that are made on the basis of scores obtained by the examinees. Typical assessments of the validity of achievement tests include expert opinions on whether the items address the skills they are supposed to measure, statistical comparisons with independent criteria used to assess the same skills, the relationship of test performance with future performance in related areas, and factor analyses.

The most critical contribution to test validity is made as the test is developed. The description of the developmental phases of CLAST suggests that the tests are developed carefully to maximize their validity. General test specifications are developed by faculty with demonstrated expertise in both content areas and testing. Faculty members also establish item specifications and write items to meet these specifications. Items are reviewed by faculty and CLASP staff for the purpose of identifying potential issues of content, measurement, and bias. Items are field tested, and data are analyzed statistically before being selected for inclusion on an operational test form.

As evidence of concurrent validity, three studies have been conducted that compare CLAST score results on four test instruments that purport to measure related skills, though none of the criterion tests is a measure of basic academic skills. Livingston (1985) compared CLAST scores with the Scholastic Aptitude Test (SAT) and the Test of Standard Written English (TSWE). Maxey (1986) conducted a similar study with the American College Test (ACT), and Llabre (1985) compared CLAST with the Florida Teacher Certification Examination (FTCE). Correlations with these instruments are summarized in Table 1.

Table 1. Summary of correlations of CLAST scores on selected instruments

	CLAST TESTS			
	<u>Reading</u>	<u>Writing</u>	<u>Computation</u>	<u>Essay</u>
SAT Verbal	.68			
SAT Math			.73	
TSWE		.59		.49
ACT English Usage		.61		.52
ACT Social Studies	.62			
ACT Math			.69	
FTCE Reading	.44			
FTCE Writing		.45		.52
FTCE Computation			.47	

The relationship with the SAT, TSWE, and ACT tests tend to be somewhat higher than might be expected, since none of these is, strictly speaking, a basic skills test. The FTCE relationships are moderate.

Two additional studies compared CLAST performance to GPA. Data for the first study were provided by the Management Information Services of the Division of Community Colleges. Data for 16,503 students were analyzed, and comparisons were made between the numbers and percentages of students that passed from 0 to 4 CLAST tests and their GPAs. The combined results of the three 1985-86 are summarized below in Table 2.

Table 2. Comparison of GPA with number of CLAST tests passed in 1985-86

<u>GPA</u>	<u>Percent Tested</u>	<u>0 Tests</u>	<u>1 Test</u>	<u>2 Tests</u>	<u>3 Tests</u>	<u>4 Tests</u>
3.5 - 4.0	15.2	0.0	0.8	0.4	4.4	94.4
3.0 - 3.4	26.8	0.2	0.9	1.3	6.1	92.4
2.5 - 2.9	33.0	0.6	1.8	2.8	9.1	85.5
2.0 - 2.4	20.7	1.3	3.2	5.0	14.6	75.9
2.0 -	3.8	1.7	3.3	6.3	15.4	73.2
Totals	99.5	0.6	1.7	2.6	9.0	86.0

As would be expected, these data show that the number of CLAST tests passed correlates positively with GPA.

In his study of CLAST standards and state university system performance, Delaino (1987) found that "approximately 79% of the students described above who would not have been permitted to transfer under the 1989 standards obtained a GPA of 2.0 or higher and approximately 24% of the students achieved a GPA of 3.0 or higher." It is interesting to note that a higher percentage of blacks (83%) in the study achieved GPAs of 2.0 or better than did either whites or Hispanics (78% each).

The content and format of the Reading test appears to be generally consistent with other reading tests that have been shown to be highly valid. The same is true for the Writing test. However, the Computation test does not yield scores that can be considered to assess what are commonly considered to be computation skills. The test measures skills in arithmetic, algebra, geometry and measurement, logical reasoning and computer technology, and statistics. Except for arithmetic, computation skills may be only required marginally for success on an item. Items measuring the other skill areas require knowledge and skills far beyond mere competency in computation. Furthermore, many of the items carry a heavy reading load, and may indeed measure reading skills as well as "computation" skills. While the test appears to be written to specifications, the title is misleading, at best. Either the name should be changed (e.g., "mathematics") to reflect the content of the test or all items that are not computation (i.e., arithmetic) should be deleted.

Reliability

Reliability indices are used to estimate the extent of the variation of scores that may be attributed by random errors of measurement. They provide an indication of the amount of score difference that might be observed were it possible to administer the test repeatedly to the same examinee. The reliability of a test is positively influenced by the number of items on the test and the heterogeneity of the group being tested. In addition, tests that are very difficult or very easy tend to have less reliability than those of medium difficulty because of lower variability in scores. The Kuder-Richardson Formula 20 (KR-20) is used to assess the reliability of the CLAST.

As a means of correcting for the influence of test standard deviation on the magnitude of the KR-20 estimates, standard errors of measurements (SEM) are also provided for CLAST tests. The SEM represents the probable standard deviation of scores for an individual who might take a large number of equivalent tests.

Table 3 presents summary raw score statistics for the CLAST tests administered in September, 1986 (0986), March, 1987 (0387), and June, 1987 (0687).

Table 3. Summary Raw Score Statistics for CLAST 0986, 0387, and 0687.

	READING			WRITING			COMPUTATION		
	<u>0986</u>	<u>0387</u>	<u>0687</u>	<u>0986</u>	<u>0387</u>	<u>0687</u>	<u>0986</u>	<u>0387</u>	<u>0687</u>
Items	36	36	36	35	35	35	56	56	56
Mean	27.64	28.26	28.44	30.66	30.69	31.05	41.05	41.57	41.20
Median	29	29	30	31	31	32	42	42	42
SD	5.65	5.19	4.81	3.56	3.36	3.15	7.62	7.19	7.23
KR-20	0.84	0.83	0.79	0.73	0.73	0.68	0.85	0.84	0.84
SEM	2.29	2.15	2.20	1.83	1.76	1.77	2.99	2.92	2.93

Based on the average mean raw score as a percentage of the maximum possible raw score, the writing test (88.0%) is the least difficult of the three, Computation (73.7%) is the most difficult, and Reading (78.1%) is in between. The reliabilities of the three tests rank in the same order, as do the SEMs. These results are consistent with the variability of the tests and with test length.

Whether these reliabilities are sufficient for the ways in which the test scores are used is largely a judgment call. The reliability indices for Reading and Computation appear to be adequate, but the reliability of the Writing test is suspect. It may be advisable to reconsider the test length in future forms, or to attempt to find items of greater difficulty.

Another factor influencing test reliability is the speededness of the test. A review of the item analyses suggests that slightly more items near the end of the test are omitted than those near the beginning and middle of the test. It is apparent that some students are handicapped because they work more slowly than others. In a study of individuals with repeated failures on the test, Greenberg and Melvin (1987) found that some of the students ran out of time. Belcher and Einspruch (1987) found that extra time resulted in a 16-point improvement for retakers of the Reading test, and 9-point increase for Writing Test retakers. Younkin (1986) found that length of time had a moderate to significant effect on CLAST scores.

Belcher (1984) examined several indices of reliability that are more relevant for criterion-referenced tests such as the CLAST. In general, her results corroborated the KR-20 analyses to a greater-to-lesser degree. King (1984) found in a test/retest study that "decision consistencies for reading, writing, and essay subtests are lower than desirable when the 1989 criteria

are used." The proportion of students consistently classified ranged from .74 to .83.

The research cited here and in the Validity section is available from CLASP.

Bias

A test is said to be biased against a particular group of examinees--or to favor a particular group--if the pattern of responses to individual items is different for that group than it is for another group. A difference in average scores between groups is not necessarily evidence of test bias. Other factors may influence such score differences, such as equivalence of exposure to the skills that are measured by the test.

All statistical analyses for test fairness are based on item-by-group interaction. Items that are proposed for inclusion on the CLAST tests are reviewed for cultural issues as well as content issues, and precautions are taken to avoid stereotyping. Typically, test developers are careful to ensure that each test is balanced with items that represent men, women, blacks, and Hispanics with equal status and, where possible, equal frequency.

Statistical analyses are performed on each CLAST test to ensure that the patterns of responses to items are similar for males and females, blacks and whites, and Hispanics and whites. Scatter diagrams are produced which plot the performance on each test item by the pair of groups under consideration. Items which deviate substantially in performance between the two groups will be graphed outside the general pattern of the remaining items. These items, called "outliers," must be further examined for content bias or other factors that may or may not cause these discrepancies.

The scatter diagrams for the nine tests under focus here were examined for outliers. The statistical procedures for generating the diagrams are objective and strictly specific. While the CLASP office has used appropriate statistical methods, the science of the identification of outliers is not as straightforward but rather somewhat subjective. In addition, a criterion diagram is often produced for two random subgroups of the same gender or racial/ethnic group to identify the degree of scatter attributable to factors other than group differences. Such a criterion was not available.

Table 4 reports the numbers of items that favored one or another group in three paired group comparisons for the 0986, 0387, and 0687 CLAST tests. For example, on the 0986 Reading test, 1/1 appears in the M/F column. This means that of the 36 items on the test, 1 item appears to favor males (M) and 1 item appears to favor females (F). This test would appear to be balanced and not favor (or be biased against) either of these groups. On the other hand, for the Black (B)/White (W) group comparison for the same test, 4 items appear to favor blacks and 8 to favor whites. While some items favor each group, more appear to favor the white group than the black group. Some would call this test slightly biased against the black group. As can be seen from the table, however, most of the tests are well balanced for group fairness.

Table 4. Numbers of items that appear to favor groups* on CLAST tests 0986, 0387, and 0687.

	READING			WRITING			COMPUTATION		
	<u>M/F</u>	<u>B/W</u>	<u>H/W</u>	<u>M/F</u>	<u>B/W</u>	<u>H/W</u>	<u>M/F</u>	<u>B/W</u>	<u>H/W</u>
0986	1/1	4/8	1/1	0/0	3/2	0/1	0/0	1/2	0/2
0387	0/1	1/0	0/1	0/1	2/0	0/0	3/2	4/1	0/0
0687	0/0	2/2	0/2	0/0	1/0	1/0	0/3	1/2	1/0

*M=male, F=female, B=black, W=white, H=Hispanic

Item Analysis

Tests are assembled from the CLAST item bank. The item bank contains the items and their statistics based on previous administrations. Each CLAST form contains 5-10 experimental items that are included for the purpose of gathering data for potential future use in a CLAST form. The experimental items are not included in the examinees' scores, but are added to the item bank.

The following statement was taken from the section on item development on page 9 of the 1985-86 CLASP Technical Report:

Items are included in regular forms of CLAST as experimental items and are not counted as scored items for students. This produces classical and Rasch item statistics for statistical evaluations of quality. Items are screened based on the following criteria: p-value greater than or equal to .40, point-biserial greater than or equal to 0.30, the percentage of examinees choosing each option, and Rasch total fit less than or equal to $1.0 + 3$ standard deviations. Item data are reviewed to identify potential dysfunctional items, not for the purpose of selecting items for the test.

Other than the statement on the Rasch criteria, which will be discussed later, the last sentence of this statement is unclear. It implies that an item may be selected for a test, even though it may be potentially dysfunctional. While most selection criteria are based on information provided by Rasch statistics, the point-biserial statistic--a measure of the correlation of correct item response with total test score--provides important information on how that item may contribute to the information provided by the test score.

The criterion that the point-biserial be greater than or equal to .30 is a worthy goal. But, as shown in Table 5, CLAST forms for 0986, 0387, and 0687 apparently did not adhere to that criterion. In practice, every test developer knows that some statistical criteria must be relaxed when an item must be selected because of more content criteria. But unless live item data were radically different from experimental item data, the 0.30 minimum for point-biserials is not a true criterion. The percentages of items with point-biserials less than 30 range from 11.1% (0986 and 0387 Reading) to 48.6% (0687 Writing).

Table 5. Point-biserial frequencies and medians for CLAST 0986, 0387, and 0687.

Range of Point Bi-serials	READING			WRITING			COMPUTATION		
	0986	0387	0687	0986	0387	0687	0986	0387	0687
0.50 +	3	2	0	0	0	0	0	1	0
0.40 - 0.49	17	16	9	4	3	2	14	9	7
0.30 - 0.39	12	14	19	24	18	15	23	26	28
0.20 - 0.29	4	4	7	7	13	17	18	19	18
0.19 -	0	0	1	0	1	1	1	1	3
Median	0.42	0.40	0.34	0.32	0.31	0.29	0.33	0.33	0.32

While it may be coincidental and not an indication of general trend, it should be noted that the proportion of items with lower point-biserial indices increases for all three tests across the three administrations. The same trend was seen, though to a lesser degree, with the KR-20 reliability estimates (see Table 3).

Item Calibration

The mechanism that makes item-bank based testing systems such as CLAST feasible is known as Item Response Theory (IRT). IRT is based on a mathematical model that facilitates the comparison of item and ability characteristics on common scales even though they are produced from different test forms administered to different groups of examinees at different points in time. IRT provides statistical estimates of from one to three item parameters--item difficulty, item discrimination (the capacity for an item to separate those who have the measured trait from those who don't), and the so-called "cheating" (or guessing) parameter (or likelihood of getting the right answer without having the trait). IRT also provides an estimate of ability, based on the number or pattern of correct responses. The process of generating these estimates is called calibration.

The CLAST system uses a one-parameter (item difficulty) version--commonly called the Rasch model--of IRT. The Rasch model is commonly used for such systems because it is much more cost-efficient to use than the more cumbersome 2- or 3-parameter models.

All IRT models, as are classical test development systems, are based on sets of assumptions. One major assumption of IRT is that the skill or trait that is measured by the test be unidimensional. One cannot, in theory, use IRT to calibrate reading items with mathematics items. This is why separate calibrations are made for the Reading, Writing, and Computation CLAST tests. The item bank for each test contains its own item scale separate from those of the other two.

The Reading test is likely to best satisfy this assumption of unidimensionality. The items are similar in construction, and there are fewer broad skill content areas addressed in Reading than in the other two areas. The items in the Writing test are more heterogeneous in format, and the skill areas that are measured are apparently more discrepant than on the Reading test. The Computation test is most likely multidimensional. The five major skill content areas represent a variety of skills not strongly related to one another. It would be useful to carefully monitor the dimensionality of all three scales using factor-analysis techniques, Yen's Q-statistic, or another commonly recognized procedure.

It is important to note here that there is some professional disagreement among psychometricians regarding the importance of the unidimensionality assumption. Some could produce data that would indicate that IRT techniques are robust, and that violation of the assumption has little effect on the characteristics of multitrait tests. Others could produce data that indicate otherwise. If they have not done so already, CLAST developers would be well-advised to examine the instruments for the robustness of the model as it is applied for the purpose of the fulfillment of the mandated CLASP goals.

Another matter worthy of comment is the method for selecting items for inclusion on the CLAST tests. In theory, IRT analyses should be performed only if the items "fit" the model. Rasch item-fit statistics are used to help select operational CLAST items. The T-test Total statistic is apparently the primary criterion for selection, although other statistics may be used as well. There is disagreement on the "best" way to determine whether an item "fits" the model that will be used.

When classical test techniques are used to develop a test, it is typical to anticipate an item overwrite factor of approximately two. That is, twice the number of items required should be written and field tested to ensure that the selected items will demonstrate desirable statistical characteristics. Items suited for tests using Rasch model techniques must satisfy even more demanding statistical criteria. Items, experimental or operational, that are found lacking in statistical quality are flagged as potentially dysfunctional. Flagged experimental items may be modified, if the nature of the difficulty seems clear, and new statistics obtained during a subsequent administration. It is not clear what happens to operational items that are flagged.

Test Equating

The problem of conserving the comparability of test scores from one test administration to another is an important one. It is essential for ensuring that the standards are the same for all examinees regardless of when they take the test. The system used for preserving the comparability of CLAST scores is excellent.

All CLAST tests in each subject area are equated using Rasch data on items stored in the item bank. Each time a test is administered, the items are calibrated on three systematic random samples of 700 records. The resulting item difficulty estimates are averaged, and these values then adjusted to the base scale established in 1982. Revised and experimental items embedded in each operational form are adjusted to the same scale. After each administration, all Rasch item data are compared to those obtained from previous administrations, and those items whose statistics vary from previous values by a predetermined amount are flagged for further examination. Item history data are maintained for six years.

Since the item difficulty statistics, expressed in logits, are all related by a common scale, new CLAST forms may be developed that are comparable in difficulty to previous forms by controlling the mean and range of item difficulty. Since these criteria must work in consort with content criteria, it may not always be possible to develop a new form that has characteristics identical to the target characteristics.

In order to control for slight differences in item parameter characteristics, the tests are equated using linking values based on differences between the 1982 administration and the current test being equated. The test is also equated to another form using similar linking procedures to provide a quality control check.

Ability estimates for each form are developed using the item difficulty logits for that form. A separate ability logit-value is calculated for each possible raw score point on the test. The ability scale is then adjusted to the metric of the 1982 administration. The score reporting metric is then generated by linearly transforming the adjusted ability scale.

Proper data management for a test development/administration/scoring project of the magnitude of CLAST is critical and complex. While the Rasch item calibration and ability estimates are computerized tasks, CLAST staff must continually monitor the data flow. The procedures for performing these tasks are set forth on a written list that requires a minimum of 75 checkpoints. Many of the steps are performed and rechecked by hand. The prescribed procedures are thorough and well conceived.

Standards

CLAST standards, expressed in terms of minimal passing scores, were established in 1984 by the State Board of Education. The passing scores were adopted as recommended by a panel of interested persons. In the judgment of this panel, the passing scores represented the minimum level of performance acceptable for the successful completion of the sophomore year in Florida's state universities and community colleges.

Table 6. Passing scores for CLAST tests.

	1984 STANDARDS			1986 STANDARDS			1989 STANDARDS		
	Scaled Score	Approx N	Raw Score %	Scaled Score	Approx N	Raw Score %	Scaled Score	Approx N	Raw Score %
Reading	260	17	47%	270	19	53%	295	25	69%
Writing	265	23	66%	270	24	69%	295	29	83%
Computation	260	24	43%	275	30	54%	295	37	66%

Table 7 reports percentages of examinees who exceeded each of the three standards since 1982-83. The data for 1982-83 through 1986-87 are for first-time examinees only. Table 8 reports the percentages of first-time examinees that passed all four tests (including the essay test) during the first four examination years.

Table 7. Percentages of examinees meeting standards by subtest.

Test	Year	Number Tested	PERCENTAGE PASSING		
			1984 Standards	1986 Standards	1989 Standards
Reading	1982-83	41,844	89.5	83.2	62.5
	1983-84	42,420	93.3	88.5	66.8
	1984-85	34,501	96.9	94.4	80.4
	1985-86	34,722	96.5	93.8	76.6
	1986-87*	35,361	--*	92.7	76.5
Writing	1982-83	41,844	93.1	89.7	59.4
	1983-84	42,420	94.6	92.6	69.6
	1984-85	34,501	96.2	94.7	74.1
	1985-86	34,722	96.7	95.2	78.8
	1986-87*	35,361	--*	95.7	79.7
Computation	1982-83	41,844	90.6	78.5	55.2
	1983-84	42,420	94.1	83.7	57.9
	1984-85	34,501	97.3	91.3	73.2
	1985-86	34,722	97.3	91.8	72.5
	1986-87*	35,361	--*	92.5	74.3

*Not computed by the CLAST Office.

Three increasing levels of standards were established. The passing scores are expressed in terms of the scaled-score metric. Table 6 reports each of the three scaled score standards, with the approximate number and percentage of items required on each test to meet the standard.

Table 8. Percentages of examinees passing all four CLAST tests.

Year	Number Tested	PERCENTAGE PASSING		
		1984 Standards	1986 Standards	1989 Standards
1982-83	41,844	72.6%	62.4%	27.5%
1983-84	42,420	80.4%	71.4%	32.3%
1984-85	34,501	88.6%	83.0%	43.7%
1985-86	34,722	88.3%	82.5%	43.7%
1986-87	35,222	--	83.9%	45.7%

The data presented in the three preceding tables provide information about differences in the level of performance that is required to meet all three standards, and about success rates over time. While the three sets of standards are increasingly demanding, the increment from 1986 to 1989 is greater than the one from 1984 to 1986. By 1989, an examinee must be successful on 83% of the Writing items in order to achieve a passing score. The cutoff point is too close to what would be a perfect score to discriminate sufficiently between those who have acceptable writing skills and those who do not. A solution to this "ceiling" problem would be to add more difficult items to the test.

Each year, increasing proportions of students have met all three standards fixed for the Writing and Computation tests. Passing proportions for the Reading test seem to have peaked during 1984-85, and have declined slightly in each succeeding year. If, as designed, the scaled scores for each subject-area test are controlled for difficulty, one could conclude that performance levels among examinees have appeared to increase in all areas since 1982, except for a decline in Reading performance for the past two years. However, if the nature of the population of examinees has changed over time, the increases may be a function of population differences. The 18.7% decline in the number of examinees tested between 1983-84 and 1984-85 may provide a clue to whether or not there have been some changes in population characteristics.

The method by which the CLAST standards were established is consistent with common standard-setting practice. The system of escalating standards anticipated and, perhaps, encouraged increasing levels of performance. The standards have been reviewed periodically against test results. This practice of review is an essential ingredient for monitoring project results with respect to project mission.

Summary

In general, the CLAST system produces carefully developed tests with sound psychometric characteristics. The tests appear to be valid measures of performance in basic skills. The tests appear to be adequately reliable, although the reliability indices for the Writing test are marginal. On the whole, the tests do not appear to suffer substantially from bias for either gender or racial/ethnic groups.

Some discrepancies were noted between stated statistical item criteria and the data generated on scored tests. In addition, some questions were raised regarding the sufficiency of item-to-model fit. It was noted that the Computation test appears to violate the Rasch assumption of unidimensionality, and recommended that the effect of this violation should be examined. The procedures for equating forms of the tests to ensure comparability of scores is excellent. And it was suggested that appropriateness of mandated standards be continually monitored using data from CLAST administrations.

CONTENT ANALYSIS OF VERBAL SECTIONS

Charge

The charge to this consultant was to review the verbal components of the CLAST to determine the appropriateness and content validity of these sections as instruments that measure the academic skills of college sophomores.

In order to address the charge, it was necessary to review:

- the history and development of the College Level Academic Skills Program;
- the history and development of the College Level Academic Skills Test;
- documents, materials, and reports on all aspects of the test;
- the test series and forms currently and previously in use;
- the test specifications;
- all test items;
- test scores and patterns of scores;
- testing procedures.

In addition, a series of interviews were conducted in order to better understand the existing policies and procedures associated with the CLAST. Finally, observations were made of key aspects of the Essay reading and scoring procedure.

The following is an in-depth review of the verbal components of the CLAST. First, the test is described briefly in an Overview and then each section is reviewed separately including a description, highlights of findings and observations according to item analyses, and recommendations for that section of the test. The sections are reviewed in the following order: Writing; Essay; Reading.

Overview

The verbal component of the College Level Academic Skills Test is intended to assess college student proficiency in writing and reading. The Writing and Reading sections are multiple choice. The Essay is an actual writing test. The total time allowed for these three verbal sections is 120 minutes.

The skills tested in the verbal component are based on specific skills called Objective Skills statements. These Statements detail what students should know to demonstrate proficiency in the area being tested.

The verbal component is a part of what may be defined as a criterion referenced test in that it sets a standard, or criteria, for performance and achievement.

Writing

The CLAST Writing test is a multiple choice test of 35 items that test 17 Objective Skills. The Objective Skills in Writing are Word Choice Skills; Sentence Structure Skills; Grammar, Spelling, Capitalization, and Punctuation Skills. Most skills are sampled an average of three times on a test, not including experimental items. Within the multiple choice format, several item forms are used. They include fill-in, alternate choice, negative response (no error; none of the above), and best answer. No specific time is allotted for the test, although the directions suggest that students use 25 minutes for the section (of the 70 minutes allowed for both the Writing and the Reading tests).

Findings

1. The items on the Writing Test are written clearly and precisely. Directions are concise. All items reflect requirements of the Specifications for Writing, the guide for those who write the test items.
2. Student scores on the multiple choice Writing Test show that the large majority of students answered almost all items correctly; particularly those items that tested grammar, spelling, capitalization, and punctuation skills. The items are written on a very basic level. The item content ranged from very simple to easy questions, and posed virtually no challenge to test takers.
3. The readability, or comparative grade level equivalent of the writing and items on the Writing Test range from what is estimated to be grade levels 7 to 10.
4. About a third of the items are in the form of alternate choice questions, where a student has to select the correct answer from a choice of two possible answers. This form increases the possibility of correct answers through guessing and may artificially inflate test scores and thus proficiency.
5. At least one item in the test series combined the item forms of "best answer" and "no error." The directions indicate the test taker is to select the correct answer based on his or her own judgment and based on fact. The two types of possible responses are not mutually exclusive.
6. While no other bias seems to exist, at least one item of vocabulary is probably not in current use in texts or in the speech of college faculty. The percentage of accurate response for this item was low. The vocabulary word, used in the context of the item, was "rout."

7. The Writing Test does not assess appropriate student writing skills. It is an objective test of isolated skills that may show how well a student knows the mechanics of English; whether a student is able to recognize correct tenses, and subject and verb agreement; whether a student is able to recognize a misspelled word.
8. Objective Skills Statements are generally listed as skills students have, or have the ability to recognize. Skills that attempt to determine whether students write fragments, fuse sentences, or use awkward constructions are listed as errors, not as skills.

Recommendations

1. The high success rate on the writing test, and the relatively low level of the test language and items suggest ease of items rather than high achievement. The standard is not reflective of what is generally considered college level. Some consideration may need to be given to increasing the difficulty level of the test or to using the Writing test as a placement instrument for entering freshmen.
2. The accuracy of the title of the test is important because it communicates to the test taker and to the test administrator, what skills and abilities the test is to measure. What is labeled as a Writing test of the CLAST does not actually measure a student's ability to write, as does the Essay. The multiple choice format more accurately measures language usage skills, and some general functions of the language.

It is recommended that the title of the Writing Test be changed to reflect the kinds of skills measured, as in "English Usage Skills," "Language Usage Skills," or the like.

3. It is recommended that some consideration be given to the process used in the development of the Objective Skills Statements. According to the Specifications for Items, spelling items are constructed based on the use of the "list of Commonly Misspelled Words." In addition, according to some interviews conducted for this report, the skills represent those most students were likely to have, and if tested, would show proficiency. It would be more reasonable and educationally sound to select spelling or other language usage items from the college English or General Education curricula than from a commercially distributed list of words for popular use.
4. Some consideration may need to be given to the overall objectives of the assessment in Writing. The main thrust of the test is not clear. The test focus could be in finding out how much and how well students are able to think and write creatively, or how well students recognize word parts, the meanings of words, or sentence and paragraph structure, or a combination of several goals. Whatever the focus, the test results should show clearly student achievement in terms of the goals.
5. Some consideration may need to be given to increasing the accuracy and reliability of the assessment of student writing. The present Writing test alone is not able to determine student writing ability. A more

accurate test would require a generalized application of the skills to a written work, such as on an Essay, or a combination of these measurements.

Essay

The Essay section of the CLAST is a separate timed test of 50 minutes, within which time students are required to "plan, write, and proofread" a written work.

Students have two topics from which to choose in the writing of the Essay. The topics require that students write of their own experiences, dealing with broad topics, such as family, education, the environment, the community. The topics require writing that is generally expository or persuasive in nature.

The Essay is scored holistically on a scale of 1 (lowest possible score) to 4 (highest possible score). Students receive a combined score from two readers.

Findings

1. The Essay of the CLAST, as an instrument that measures writing skill, is appropriate.
2. The method of holistic scoring for assessing student writing skill is sound, reasonable, and appropriate.
3. The Essay reading and scoring procedure is thorough, efficient, and comprehensive. A potential problem of unreliability with readers of divergent objectives and standards is effectively addressed by frequent review of the scoring standards and ranges before and during a scoring session.
4. Statistical data on the essay, including interreader reliability, percent of agreement in scoring, individual reader analyses (coded), and similar information all appear to be within normal limits.
5. The method of equating the essay scores from administration to administration currently used in the CLAST program is inadequate to assure appropriate comparability of scores. As an essay topic changes, so does its difficulty and the way it is perceived by the writer (i.e., the student). The only acceptable psychometric method of equating essay scores is to relate them to an anchor set of multiple choice items.
6. The 4-point scale on which Essay scores are currently based severely limits the range of possibilities for scoring. Analysis of the results demonstrates that only 9 to 10% of the papers were given a score of "4" (the highest possible score) by the readers, while only 6 to 7% received a "1" (the lowest possible score). This probably reflects a combination of the actual level of the papers and the tendency of some readers to avoid the riskier extreme scores.

7. The directions for writing the Essay are clear and give students effective cues for planning, getting started, and developing the Essay.
8. Timing for the Essay is adequate. The Essay requires no new information, and the prompt is precise. Planning time is assisted through the directions. Re-take statistics show no appreciable difference in scores with additional time. Results show students are generally able to complete the writing of the Essay.
9. The standards used for grading and scoring Essays of the CLAST would probably be more appropriate for student placement into entry level or remedial courses. More rigorous standards may need to be considered in assessing students who have completed two years of college-level study. College sophomores are generally expected to combine successfully elements of rhetoric, grammar, and style in presenting information of their own and that of others. Those expectations are not apparent in the current practice of Essay grading.

RECOMMENDATIONS

1. It is recommended that some consideration be given to the following:
 - that writing skills be taught throughout the curriculum in the system;
 - that a writing assessment on the sophomore level require that a student show the ability to think critically and objectively in presenting information; write for a specific purpose, as to persuade, describe, analyze, inform; present information in any specific and different tone, mood, or purpose; synthesize information and present in a logical, unified, and cohesive manner; interpret and evaluate information; extend an argument and create one's own argument; express the ideas, thoughts, and feelings of others; understand and use language accurately, creatively, efficiently, and effectively; understand the mechanics and uses of language.
2. Combine the Essay with the Writing section. This will not only improve the equating procedure but also provide a more accurate picture of students' writing.
3. Consider increasing the holistic scoring scale to six points. This would allow for greater differentiation of scores and increase the accuracy of measurement.

Reading

The CLAST Reading test is a multiple choice comprehension test with an average of 36 questions, not including experimental items. Twelve Objective Skills are tested. The Objective Skills are broadly classified as Literal Comprehension Skills and Critical Comprehension Skills. Each specific skill is sampled an average of three times per test.

Within the multiple choice format, several forms are used, including fill-in, alternate choice, and negative response. No specific amount of time is allotted for the test, although the directions suggest that students use 45 minutes for the section.

Findings

1. The CLAST Reading Test has performance standards set at the level of minimum basic skills. The reading passages vary in reading grade level equivalent from grades 7-12, somewhat lower than the level of most college-level textbooks. There is, however, no significant difference in the frequency of correct responses by the students at any grade level. Student success rate is consistently high, particularly with Literal Comprehension Skills questions.
2. The Objective Skills Statements for Reading are minimum basic skills. The three Literal Comprehension Skills require only that a student recall what was read and figure out a vocabulary word in the context of the passage. The skills do not require any manipulation of information or thinking skill that would be developed as a result of the modes of inquiry associated with college-level instruction. The Critical Comprehension Skills are more complex and require more student participation in producing the answer, but require no other skill or information than what appears in the passage. They are very close to Literal Skills in requirement.
3. Literal Comprehension Skills and Critical Comprehension skills are tested at the same frequency throughout the test series, without a particular emphasis or focus on what may be essential skills that a college sophomore would need to have and for which to demonstrate proficiency.
4. Passage questions are generally precise and clear in intent and meaning. Some, however, include distractors, or possible answers, that are either very close in meaning to the correct answer, or are so distant that it would be nearly impossible to be the correct answer. On at least two occasions, the correct answer was much longer, and more clearly stated than other distractors. This provides clues to students who may simply be good or practiced test takers, rather than students who actually know the answers. Some examples of weak distractors include: Test 9/86-page 1, question 1D; passage 2, question 6 A-D; passage 3, question 15 A-D, question 16 A-3.
5. Passage questions clearly indicate for students the type of question and the kind of information required to answer each question correctly. For example, the item stems state the kind of response required, such as main idea, relationship between sentences, tone, intent, mood, valid argument. While this offers the advantage to students of not being misleading, it adds to the ease of the items and so to the ease of the test. Written in the format of explicit statements and strong cues, the item responses do not require that students be able to recognize or determine the actual skills being tested. Rather, the items require that students be able to show examples of the skills as stated in the stem. In addition, the items as written make it possible for students to study for the test, and allow others to teach to the test. A test taker may predict that each skill will be tested an average of four times per test and will be reasonably assured of the kinds of questions asked. It then becomes more difficult to distinguish how much a student knows from how well s/he is able to take the test.

6. The passages are generally precise and clear in presenting information. Some passages, however, have unnecessary words and phrases, with generous use of commas, connectors, nonfunctional words, and the like. These types of constructions add to the length, not necessarily to the difficulty of the passage.
7. Test questions appeared in three forms of multiple choice format: fill-in, negative response, and alternate choice. There is only one fill-in per test, and an average of 3 of each of the other two formats. This lack of variety in item format lends to teaching to the test and promotes ease of items.
8. The timing for the test is adequate, particularly when viewed in terms of the effect of time in re-take situations. However, students are expected to be self-paced, which offers the possibility of the misuse of time for the completion of two distinct tests, and which may affect some students' scores.

Recommendations

1. It is recommended that the standard of performance on the Reading test of the CLAST be raised to reflect college-level performance standards.
2. It is recommended that the current CLAST Reading test be considered for use as a placement test at entry level for placement into remedial college courses.
3. It is recommended that the Objective Skills in Reading be reviewed. It is further recommended that some consideration be given to the following:
 - that critical reading skills be stressed throughout the curriculum;
 - that a reading test at the college sophomore level require that a student show the ability to think critically, interpret and evaluate information; understand abstract material; understand relatively complex theories and concepts; analyze and appraise written work; draw conclusions and make sophisticated inferences; make informed judgments; summarize and generalize written material; perceive close distinctions of thought; perceive nuances in meaning; distinguish between author's thought on a subject and that of the reader; perceive tone, mood, and purpose of the author; understand and use facts and concepts and apply to other situations; read efficiently and with speed; and understand the different purposes for reading; such as previewing, finding bits of information, and skimming. These skills move beyond the level of basic skills into college-level expectations. They reflect elements of most college-level textbooks.
4. Some consideration may need to be given to the objectives of the reading test. The results of a reading test, depending on its focus, could provide information that would assist in the diagnosis of student strengths and weaknesses; be useful in helping to design or re-design courses, or otherwise influence the curriculum in all areas. The current test, as constructed, would not prove useful in providing those kinds of information about the instructional process.

5. It is recommended that more attention be given to the screening of items for the test, particularly in view of the high success rate experienced even with experimental items. Reading passages and questions may be made to be more varied, and less predictable.

CONTENT ANALYSIS OF COMPUTATION SECTION

I. Foreword

A. Intent of the College-Level Academic Skills Test

From: CLASP MATERIALS FOR REVIEWING CLAST STANDARDS,
Fall, 1987; pg. 10

Purpose of CLAST

"The College-Level Academic Skills Test has been designed as an achievement test. It is intended to measure the level of achievement of the communication and computation skills which are expected of all students by the time they complete their sophomore year in college, i.e., those listed in State Board of Education Rule 6A-10.031, FAC."...

From: CLASP TECHNICAL REPORT, 1985-86; pg. 5

DEVELOPMENT OF CLAST, Skills Identification

"The state-level task forces, together with the project director and other project personnel acting in an advisory capacity, worked to identify essential academic skills which every student, regardless of major, should have acquired by the end of the sophomore year." ...

Note: Emphasis (underlining) added by the reviewers,
C. Pine and R. Voronka

B. Guidelines Used by Reviewers

The criteria we used in our examination of the mathematics subtest for validity of content stem from the statements concerning CLAST which were quoted in the preceding. The expectation that all students pass the computation test by the end of their sophomore year in college implies that the mathematics skills involved should be those that are essential for successful performance in junior and senior level college courses and, beyond that, essential for the college graduate as an educated person. Further, that these skills be essential for every student regardless of major clearly imposes the restriction that the skills be neither program specific nor

course specific. Thus for example, specialized skills or specialized knowledge which might be considered useful, or even essential, for say a student majoring in one of the physical sciences should not necessarily be tested on the CLAST. Similarly, skills or terminology which might be considered necessary for a particular course, say a calculus course, should not for that reason only be tested on CLAST unless that course could be identified as a course which should be required of all students regardless of major.

We think that the four generic competencies identified: algorithms, concepts, generalizations, and problem solving are appropriate for the test. After reviewing the broad-skill categories in terms of which the competencies are tested, we examined the specific skills, specifications, and items with regard to the following questions:

1. Are the skills essential, are they reasonable to expect, and are they sufficient?
2. Do the item specifications describe the skills appropriately?
3. Do the items conform to the item specifications?
4. Is the test appropriately balanced in content?

Our judgments with regard to these questions are necessarily subjective ones, and stem from our longtime teaching experiences and our service on the mathematics Advisory Committee to the New Jersey Basic Skills Council. The context in which these judgments were made was that of the validity of the Computation test as a whole in serving its intended purpose: to provide a reliable assessment of achievement in mathematics necessary and reasonable for students completing their second year in college.

The materials available to us for this evaluation included:

- the set of item specifications applicable to the 1984 code
- the set of item specifications applicable to the 1987 code
- three forms of the test coming under the 1984 code: Sept., 1986; March, 1987; and June, 1987
- one form of the test coming under the 1987 code: Sept., 1987
- proportion of students choosing the various options for each of the items in the Sept., 1986; March, 1987; June, 1987 but not the Sept., 1987 tests.

II. Comments and Recommendations

As a general comment, we feel that the computation test has been improved considerably by the revision from the 1984 code to the 1987 code. This marked improvement necessitates fewer recommendations for change than we would otherwise have made. Also, we want to remark upon the thoroughness of the item specifications. Obviously a great deal of time and thought went into writing of these specifications.

Our major reservation about the test lies with the broad-skill category Logical Reasoning and Computer Technology. We agree heartily with the deletion of the computer technology skills in the 1987 code. However,

we would go beyond this. Our major recommendation is that the Logical Reasoning category be completely eliminated. While we recognize that the ability to reason logically is essential, we feel that most items generated by the pertinent specifications should not be a part of the mathematics test. Indeed, in our judgment, the inclusion of such items tends to obscure the meaning of the total test score. Beyond that, we would not want such topics taught in high school and lower-division college mathematics classes at the expense of essential mathematics skills. The recommendations and comments which follow pertain to the skills, specifications, and items included in the broad-skill categories arithmetic, Geometry and Measurement, Algebra, and Statistics, including Probability. Our frame of reference will be the test as it exists at present for these four broad-skill categories (that is, the skills and specifications coming under the 1987 code) and we will refer to the skills and items, including those on the three earlier forms of the test, by the 1987 code designations.

A. Recommended Deletions of Skills

	<u>Skill #</u>	<u>Category</u>	<u>1987</u>	<u>Skill Statement</u>
1.	20	Geometry and Measurement	2B2	Classifies simple plane figures by recognizing their properties
2.	23	Algebra	2C1	Recognizes and uses properties of operations
3.	40	Arithmetic	3A2	Selects applicable properties for performing arithmetic calculations
4.	44	Algebra	3C2	Selects applicable properties for solving equations and inequalities
5.	54	Statistics	4D1	Solves real-world problems involving the normal curve

Discussion of these recommendations:

1. Skill 20 Each of the items measuring this skill in the four test surveyed involved merely knowing the meaning of one of the following words: acute, obtuse, or right. This does not seem appropriate for "demonstrating understanding of geometric concepts." We suggest that this skill be eliminated if better item types can't be devised.
2. Skill 23 In our view, one of the contributing factors to the difficulty many students have in learning mathematics lies in what this skill involves with regard to Format 1 and Format 2. (We do not object to Format 3.) It relates to the "abstract to concrete" approach to teaching mathematics rather than the "concrete to abstract" approach that we would advocate. Students, of

course, need to understand the associative, commutative, and distributive concepts but in an intuitive way and not necessarily by name. We would want students to understand that $2(3+4)$ is equivalent to $2(3)+2(4)$ because that is what will yield 14 as the result. They have verified this rule repeatedly and accepted it intuitively. Our main concern here is the effect that the presence of this skill on the test may have on the curriculum. We don't want to encourage the current practice in many courses of having students work on remembering definitions such as " $a(b+c)=ab+ac$ is called the distributive property of multiplication over addition". It is most unfortunate that so many students come away thinking that mathematics consists of memorizing words, rules and processes.

3. Skill 40 The preceding discussion of Skill 23 applies here also.
4. Skill 44 Our concern here is similar in spirit to our concern about Skills 23 and 40. If, for example, we want to test whether students really understand how to solve linear equations, then we would have them solve some number of equations of varying degrees of difficulty rather than discuss in an abstract way the solution of a simple linear equation. Again, our primary concern here is the effect that the presence of this skill on the test can have on curriculum, in encouraging an abstract approach to the teaching of mathematics.
5. Skill 54 In our judgment, knowledge of the properties of a normal distribution is rather specialized to be considered knowledge which all students should retain.

B. Recommended Revisions of Skill Specifications

<u>Skill #</u>	<u>Category</u>	<u>1987 Code</u>	<u>Skills Statement</u>
1. 1	Arithmetic	1A1a	Adds and subtracts rational numbers
		1A1b	Multiplies and divides rational numbers
2	Arithmetic	1A2a	Adds and subtracts rational numbers in decimal form
2. 8	Algebra	1C4	Solves linear equations and inequalities
3. 26	Algebra	2C4	Recognizes regions of the coordinate plane which correspond to specific conditions
4. 29	Probability	2D3	Identifies the probability of a specific outcome in an experiment
55	Probability	4D2	Solves real-world problems involving probabilities

Discussion of these recommendations:

1. This is not a major point. It concerns the revision made in going from the 1984 code to the 1987 code: the change from "positive rational numbers" to "rational numbers." This puts together two kinds of skills. Our preference, partly with regard to possible item analysis for curriculum improvement diagnostic purposes, is to change back to "positive rational numbers" and to add a separate skill involving operations with signed numbers. Along with this we want to convey our feeling that there is not quite enough work with fractions on the test. We think that an item involving the simplification of complex fractions would be informative.
2. We recommend strongly the upgrading of the specifications for Skill 8 to include the solution of linear equations with fractional coefficients and with literal coefficients as well as with integer coefficients. In our view the ability to solve linear equations is a most important algorithmic skill. We suggest including on each form of the test a separate item for each of the above mentioned types. We recommend along with this that Skill 8 refer only to the solution of linear equations and that a separate skill be added for the solutions of linear inequalities.
3. We suggest modifying the specifications so as to restrict the shaded region to regions bounded by lines parallel to the coordinate axes; for example, as in Item 9 (Sept., 1987). Requiring students to identify regions defined by inequalities other than the simplest is perhaps too demanding.
4. The same comments apply to both skill 29 and Skill 55. As written, the specifications allow for the possibility of problems which are too complex and difficult beyond what should reasonably be expected of all students.
 - a. We recommend that B.3 for Skill 29 be changed to read "Sample spaces should be discrete and obtainable with ease from the Fundamental Principle of Counting and contain no more than 50 outcomes."
 - b. For B.5, Skill 55 we recommend that it be modified to read "Sample spaces should be discrete and obtainable with ease from the Fundamental Principle of Counting. Numbers should be positive rationals, with integers less than or equal to 100 and with numerators and denominators of reduced fractions less than or equal to 30."

C. Recommended Additions of Skills

1. The ability to deal with percent is a skill which is essential for everyone. We think it good that the revised test has moved in the direction of greater emphasis on this topic through the addition of Skill 57 (Calculate percent increase and percent decrease) and Skill 62 (Solves real-world problems which do not require the use of variables and which do require the use of percent). However, we feel that certain other percent items are needed.

Consider the following examples of the three basic variations of the percent algorithm:

- a. What is 40 percent of 200?
- b. 80 is what percent of 200?
- c. 80 is 40 percent of what number?

We have found in our experience in testing in New Jersey that many students who can answer question a. correctly may still have difficulty with questions b. and c.. Our recommendation is to add this basic percent calculation as an algorithmic skill, perhaps using two of the three variations on each form of the test. We suggest further that since percent increase and percent decrease questions lend themselves well to real-world problems, Skill 57 be deleted as a separate algorithmic skill and instead be incorporated in Skill 62. We would encourage the continued use of other types of real-world problems involving percent. For example, Item 60 (June, 1987) struck us as a particularly good percent item requiring reasoning.

2. We recommend the addition of a skill dealing with the solution of simultaneous linear equations in two variables. This situation occurs frequently in problem solving.
3. The ability to translate a verbal expression into symbolic form, that is to be able to go from the expression "three more than twice the number N" to " $2N+3$ " is a subskill of the ability to translate verbal relationships into equations. Lack of this latter ability contributes greatly to the difficulty so many students have with word problems. We recommend that skill be added to test this essential ability, both on the expression level and on the equation level.
4. We want to note our wholehearted agreement with the addition of Skill 61 in the revision of the test from the 1984 code to the 1987 code. The ability to estimate is a most useful and necessary skill and certainly should be tested. We like the item type used: Item 54 (Sept., 1987).

III. Conclusions

The additions and deletions of skills made in the revision of the test from the 1984 code to the 1987 code are laudable. We suggest the continuation of these changes in the spirit of those made and have provided recommendations in this report.

We commented earlier on the thoroughness of the item specifications. The item specifications do describe the skills appropriately.

We found that the items conformed well to the item specifications. In a few cases, where the items for the same skill appeared somewhat different on the four tests, upon review of the specifications we concluded that such diversity was allowable.

An important consideration in the evaluation of the test is the question of content balance, i.e., whether the item distribution among the skill categories is in appropriate proportion to the relative importance of the categories. For the population for whom the test is intended, in our judgment algebra and arithmetic are the most essential of the broad-skill categories. The content balance on the Sept., 1987 test clearly has shifted in this direction, and we indeed concur with this change. The recommendations we have made in this report support this trend.

CLOSING REMARKS

The State of Florida has made a significant contribution to assessment in higher education with the implementation of the College Level Academic Skills Test. Its efforts should be applauded, particularly since the endeavor moves into uncharted areas of performance-based testing on a statewide level.

A part of the on-going concern over the performance standards as established by the state has to do with the impact the current and proposed standards may have on particular segments of the higher education population. It is apparent that those Black, Hispanic, Native American, and English as a Second Language students who do take the CLAST consistently do not perform well.

While certain groups do not perform as required, it is crucial that the State remain resolute with the application of the standards. Simultaneously, the State should provide maximum opportunities for all students to perform successfully on the test, and ensure equity in the process. The State should consider a host of academic intervention activities to this end. Such intervention would necessarily include basic skills and developmental education programs, counseling classes and programs, comprehensive learning assistance and tutoring programs, and courses and classes designed to improve basic skills.

Students should have these programs made available at entry into college, before taking college-level courses and the CLAST, and having to experience failure and a disruption of studies.

A well designed and maintained basic skills program would provide a solid academic foundation for all students, but particularly for those who need to overcome histories of educational disadvantage. A program of this nature would ensure that all students have the opportunity to achieve at minimum standards, but, more importantly, at the standard and level of achievement representative of college level.

As well, the State needs to give serious consideration to the notion of teacher accountability for student success at all levels of the educational system, and move to provide methods for development and implementation of such.

The challenge facing Florida, and other states as well, is in providing both a quality and an equitable system of education. The establishment of standards and a program of assessment is an important first step.

Appendix A

LIST OF PERSONS INTERVIEWED

JAKE BEARD
Technical Advisor to CLAST
Florida State University

MYRON BLEE
Former Director, CLAST

TOM FISHER
Director of Testing
Florida State Department of Education

JAMES HELMS
Staff, Florida House of Representatives
Higher Education Committee

SYLVIA HOLLADAY
Chair, CLAST Communication Task Force
St. Petersburg Community College

TERRY HATCH
Staff, Florida House of Representatives
Higher Education Committee

DAN KELLY
Chief Reader
CLAST Essay Exam
University of Florida

SUSAN LEGG
Associate Director for Testing and Evaluation
University of Florida

JOHN McLANAHAN
Test Administrator
Florida State University

MARSHALL MILLER
Contractor, Computation Section
Tallahassee Community College

ELIZABETH NOVINGER
Associate Chief Reader
Tallahassee Community College

ADRIENNE PERRY
Chair, CLAST Task Force
Seminole Community College

ERNEST ROSS
Former Chair, CLAST Task Force

ROBERT STAKENAS
Chair, CLAST Standing Committee
on Student Achievement
Florida State University
(Interview conducted by phone)

STAFFORD THOMPSON
Legislative Analyst
Senate Education Committee

RALPH TURLINGTON
Former Florida Commissioner of
Education 1974-1986

EVA WANTON
Dean of General Studies
Florida A&M University

JEANINNE WEBB
Director of Instructional Resources
University of Florida

WILLA WOOLCOT
Chief Reader, CLAST
St. Petersburg Community College

Appendix B

The Contractors

Dr. Edward A. Morante was the lead consultant. He is a psychologist with extensive experience in testing and assessment. He has worked in higher education for more than 17 years as a college counselor, administrator, and statewide coordinator. From 1981 to 1984, he was Director of the New Jersey Basic Skills Assessment Program and is currently the Director of the College Outcomes Evaluation Program for the New Jersey Department of Higher Education.

Dr. David Osberg is the Director of Research and Measurement at The Riverside Publishing Company in Chicago, Illinois. His primary responsibility is to design, implement, execute, analyze, and interpret the research necessary to develop, maintain, and improve numerous standardized tests. Dr. Osberg had had extensive experience in a wide variety of classical and ICT methods, and works closely with test users in helping them to apply test results in appropriate fashion.

Ms. Nina Jemmott is currently the Coordinator for the Consortium for Pre-College Education in Newark (New Jersey). She has been involved with teaching reading and writing and administering college programs in developmental education for more than a decade. Ms. Jemmott has served on the Student Learning Outcomes Committee developing a statewide test for New Jersey college sophomores and is currently a member of both the College Outcomes Evaluation Program Council and the Basic Skills Council in her state.

Dr. Charles Pine is Professor Emeritus of Physics at Rutgers University where he taught physics and mathematics for almost 40 years. Winner of four outstanding teacher awards at Rutgers, he was named National Professor of the Year in 1984. Dr. Pine was a member of the New Jersey Basic Skills Council for nine years and Chaired its Mathematics Advisory Committee which developed the math sections of the New Jersey College Basic Skills Placement Test. His current focus is the training of high school and college faculty in the teaching of mathematics.

Dr. Roman Voronka is Professor of Mathematics at New Jersey Institute of Technology (Newark) where he has taught numerous mathematics courses over the past twenty-five years. Last year he was named teacher of the year at NJIT. Dr. Voronka has been actively involved in writing items and designing mathematics tests both at his own institution and as a member of the Mathematics Advisory Committee in New Jersey.